WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal Homepage: www.ijrpr.com ISSN 2582-7421

A Review on the Cooling Techniques Implemented in Data Center Cooling

V. Prasuna[1], Y. Meghanadh[2], V. Nivash[3], Vijaybabu[4]

[1-4] Department of Mechanical Engineering, GMR Institute of Technology, Rajam-532127, Andhra Pradesh, India
Email: Prasuna.vavilapalli@gmail.com

ABSTRACT

Rising computing density is increasing the thermal and energy-management demands placed on data centers. This review synthesizes eight supplied publications in the requested sequence: helium-based air-side cooling; two studies of transient spray cooling and quenching; a review of data-center cooling systems; direct air-side free cooling; multi-zone model-predictive control; rear-door heat exchangers (RDHx); and nanofluid-assisted immersion cooling. The papers span experimental work, analytical modeling, computational fluid dynamics, system simulation, and review research. The helium study reports a 1.45-times heat-exchanger conductance relative to air and lower modeled cooling energy in five Chinese climates. Spray studies clarify regime transitions and show that pressure and droplet-scale behavior can materially affect heat removal, while also exposing limits in current predictive inputs. At facility scale, optimized free cooling and zone-aware control reduce modeled energy or cooling supply, whereas RDHx and nanofluid-assisted designs illustrate higher-density heat-removal options with important infrastructure and pumping trade-offs. The evidence does not support a single universal cooling winner: performance depends on scale, climate, heat load, operating constraints, and the metric used. The review identifies validation, reliability, and lifecycle economics as priorities for future work.

Keywords: data-center cooling; helium; spray cooling; quenching; free cooling; model predictive control; rear-door heat exchanger; immersion cooling; energy efficiency

1. INTRODUCTION

Data centers convert electrical power into computation and heat. As rack and chip power rise, cooling must protect equipment temperatures while limiting the electricity, water, and maintenance burden of heat rejection. Air cooling remains attractive because it is mature and familiar, but its performance depends strongly on airflow delivery, containment, and the temperature rise permitted across the equipment. More intensive liquid-based or rack-level approaches can remove heat closer to its source, but may require additional pumps, distribution hardware, controls, and operational expertise [4,7,8].

The eight selected papers address this problem at different scales. One experimental study tests helium as a replacement gas in a sealed modular data-center cooling loop [1]. Two spray-cooling studies examine heat transfer at hot surfaces: one develops a physics-based transient model across boiling regimes [2], while the other experimentally examines quenching under varying system pressure [3]. A broad review organizes data-center cooling improvements across air distribution, water temperatures, equipment, and intelligent management [4]. The remaining studies examine direct air-side economizer operation [5], multi-zone CRAC control [6], rack-level RDHx [7], and a nanofluid-assisted server cooling plate within an immersion architecture [8].

The purpose of this review is to compare the mechanisms, reported outcomes, and limitations of these papers without treating unlike tests as equivalent. In particular, spray cooling can inform high-heat-flux thermal management, but the supplied spray studies do not demonstrate deployment in an operating data center. Similarly, modeled savings and component-level gains should not be interpreted as directly interchangeable facility-wide energy savings.

2. REVIEW METHODOLOGY

This narrative review is based on the eight supplied publications, read in the sequence requested. The set includes experimental, numerical, analytical, and review studies. For each paper, the cooling scale, method, working fluid or control variable, principal quantitative result, and stated or inferable limitations were extracted. Findings were then grouped by scale: substrate/component, server/rack, and facility/system. The review is not a systematic database search and does not claim exhaustive coverage of the field. Quantitative results are attributed to the individual study and retained with their original metric and benchmark; cross-paper comparisons are descriptive rather than a normalized technology ranking.

3. REVIEW OF SELECTED STUDIES

3.1 Helium-cooled modular data center

Wang et al. experimentally compared helium and air in a sealed, finned-tube heat-exchanger test loop designed to represent a data-center rack and backplane cooling arrangement [1]. At 30 °C, the paper reports helium thermal conductivity about 5.9 times that of air and specific heat about 5.16 times as high, while helium density is about 0.14 times that of air. The measured overall conductance (KA) was 303.5 W/K with helium versus 209.0 W/K with air, a ratio of 1.45. Under equal fan power, the modeled single-rack heat dissipation improved by 13–39%, reaching 36.7 kW in the tested operating range. A system analysis across five climates estimated 17.4% lower cooling-system energy in Beijing for the helium case.

The attraction is a potential retrofit path that preserves much of the air-cooled server architecture while improving gas-side heat transfer. The trade-off is that helium's lower volumetric heat capacity can constrain heat transport at equal flow, and the concept depends on reliable gas containment. The authors identify sealing, leakage, maintenance in a sealed environment, and helium supply economics as deployment challenges. The experiments are a modular test-bench demonstration, not a long-term commercial data-center trial.

3.2 Transient spray-cooling prediction

Banerjee, Tropea, and Seshadri present an analytical and numerical method for predicting transient spray cooling as a substrate moves from film boiling through transitional boiling to nucleate boiling [2]. Instead of relying only on a bulk heat-transfer coefficient, the model estimates heat removed by individual droplets using regime-specific expressions, sums contributions using the local droplet impact number flux, and solves transient conduction in the substrate with a finite-difference method. The approach is validated against published experimental temperature histories and used to explore droplet size, velocity, mass flow, and nozzle distance.

The computations indicate that increasing spray mass flow generally accelerates cooling, and smaller droplets can improve heat removal at a fixed mass flow by increasing the number of impacts and reducing drop-drop overlap. The model also shows that velocity effects depend on regime: higher impact speed can increase spreading but shorten contact time, so the net film-boiling response is not simply monotonic. A key strength is extensibility to varied distributions and geometries; a key limitation is the need for local droplet statistics that many experiments do not report. The paper also notes that its Poisson-based correction for drop interactions has limited direct experimental validation.

3.3 Pressure and quenching in spray cooling

Kita et al. performed water-spray tests on polished iron disks heated to approximately 730 °C, with system pressure varied from 0.1 to 0.5 MPa in an argon environment intended to limit oxidation [3]. The measured onset temperature for quenching increased as pressure increased. A mechanistic model based on force balance at the droplet-vapor interface reproduced the pressure trend more closely than the other models assessed. At a volumetric flux of 54 L/(m²·min), the average total heat-transfer rate during film boiling increased by approximately 105% between 0.1 and 0.5 MPa; the sprayed water accounted for more than 80% of total heat transfer in that regime.

The analysis links higher pressure to a higher saturation temperature and a thinner vapor film, increasing heat transfer. Estimated heat removed per droplet lay between idealized Leidenfrost-film and direct-contact limits, supporting the authors' hypothesis that partial liquid-solid contact may occur before the apparent quenching point. The experiment clarifies mechanisms under controlled conditions, but its polished iron surface and oxygen-limited environment do not reproduce the variable oxide scales, roughness, geometry, and spray distributions found in industrial quenching or electronic assemblies.

3.4 System-wide data-center cooling strategies

Zhou, Gu, and Ma review data-center cooling research across four connected areas: air distribution, chilled-water supply and return temperatures, equipment and system configurations, and intelligent management [4]. Their discussion emphasizes that raised-floor airflow, aisle containment, baffles, and higher chilled-water temperatures can improve thermal conditions or chiller efficiency, but outcomes depend on the facility layout and IT operating envelope. They also describe the complementary roles of ambient/free cooling and mechanical refrigeration: free cooling can reduce chiller workload when weather permits, while mechanical cooling remains necessary under unfavorable ambient conditions or as backup. This review provides a useful system taxonomy rather than one directly comparable experiment; its reported examples are drawn from diverse sites, models, and equipment.

3.5 Direct air-side free cooling and holistic optimization

Zhou et al. develop a cross-system model and optimization framework for data centers using a direct air-side economizer (DASE) [5]. The model includes the cost of moving outdoor air, supply-air temperature and humidity, outdoor weather variation, workload, server temperature, and power-distribution losses. A holistic optimization then selects operating setpoints across IT, cooling, and electrical systems. The authors report that the optimal free-cooling cost curve is piecewise, that supply-air fans can account for more than 92% of cooling energy in some conditions, and that optimized operation can save over 5% of total energy during summer weeks and over 46% of cooling energy during winter weeks. The denominators and seasons differ, so these percentages should not be compared as if they were the same metric.

The study's main contribution is its explicit accounting for the fan and conditioning costs that accompany outdoor-air use, rather than treating free cooling as energy-free. The framework also identifies four operating modes across changes in weather and workload. Its evidence is simulation-based for a modeled facility; site-specific filtration, humidity, air quality, weather, and control reliability remain important when applying the results.

3.6 Multi-zone CRAC model-predictive control

Zhao et al. propose cooling-supply allocation model-predictive control (CSA-MPC) for a room-level CRAC system divided into five virtual thermal zones [6]. An attention-based prediction model estimates zone temperatures while accounting for neighboring-zone thermal coupling. Particle-swarm optimization selects the supply-air temperature and airflow for each CRAC unit, with measured temperatures fed back into the rolling control process. In the selected case study, the maximum inter-zone temperature difference decreased from 7.05 °C under baseline control to 1.02 °C under CSA-MPC, a 6.03 °C reduction. The room was controlled around 28 ± 0.5 °C, and reported total cooling supply fell by 8.73% relative to baseline.

The result illustrates that better allocation can reduce both overcooling and temperature imbalance without changing the room layout. The paper reports model validation against field measurements and simulation, but the quantitative control comparison centers on a particular five-zone room and representative operating period. Transfer to other facilities requires retraining or recalibration, and operational deployment should examine sensor faults, changing workloads, optimization latency, and safe fallback control.

3.7 Rack-based rear-door heat exchanger

Cho evaluates a rack-based RDHx system combining air-cooled servers with water-cooled heat exchangers mounted at the rear of the racks [7]. The design analysis describes a 150 kW module comprising ten 15 kW racks, with supply air at 24 °C and water supplied to the secondary cooling distribution unit at 14 °C; the central chiller supply is modeled at 12 °C. CFD results indicate nearly uniform rack inlet temperatures of about 24.12–24.29 °C for the ten-rack module. A separate annual energy model for a 30 MW reference data center gives a cooling PUE range of 1.25–1.33 and an average of 1.273. In that modeled energy use, chillers contribute about 46%, fans 21%, and pumps 33% of rack-cooling energy.

RDHx shortens the air path and captures heat at the rack, potentially enabling warmer chilled water and reducing air mixing. The result remains a combination of CFD and energy modeling rather than a completed field demonstration. The paper also describes the module as 150 kW in its design and methods but refers to a 200 kW module in its conclusion; the 150 kW value is used here because it matches ten racks at 15 kW each. This internal inconsistency should be resolved before using the conclusion's module rating for design.

3.8 Nanofluids in an immersion-cooling architecture

Alkasmoul et al. use CFD and conjugate heat-transfer simulation to assess Al_2O_3-water, TiO_2-water, and CuO-water nanofluids in the cooling plate of a server model, within an enclosure containing dielectric HFE liquid [8]. Thus, the study does not immerse the server in a water-based nanofluid; the nanofluid circulates through a serpentine cooling plate while the electronics are surrounded by dielectric liquid. Across Reynolds numbers of 200–700 and nanoparticle volume fractions up to 4%, higher concentration generally lowers chip temperature but increases pressure drop. For Al_2O_3-water, the paper reports an approximately 3 °C difference between 0 and 4 vol.% at Re = 200, narrowing to about 1.25 °C at Re = 700; pressure drop at 4 vol.% is reported to rise by about 600% relative to water in the tested case.

The economic model finds pure water at Re = 600 to be the lowest-cost option under its assumed electricity price, pump efficiency, system lifetime, and fluid cost inputs. The reported minimum is $706.98 in the results and conclusion, while the abstract gives $706.15; the paper therefore contains an unresolved cost discrepancy. The analysis also assumes a small modeled server/CPU system and omits some practical issues such as long-term dispersion stability, corrosion, and maintenance. Its central lesson is that thermal improvement alone does not establish economic advantage.

4. COMPARATIVE RESULTS AND DISCUSSION

Fig. 1. Cooling concepts grouped by the scale addressed in the eight selected papers. The spray studies inform substrate-level physics; they do not report data-center deployment tests.

Figure 1 shows why these papers should be read as complementary rather than as head-to-head trials. The spray studies address the interface between droplets, vapor, and a hot solid. The helium, RDHx, and nanofluid studies address different server or rack heat-removal arrangements. The free-cooling and MPC papers focus on control and facility operation. Improvements at one scale can be limited by another: better heat-exchanger conductance does not automatically reduce total facility energy if fan, pump, chiller, or containment losses dominate.

4.1 Summary of methods and evidence

Study	Method / scale	Reported finding	Main limitation
[1] Helium	Sealed experimental loop + climate/system model	KA = 1.45× air; up to 39% more rack heat at equal fan power; Beijing cooling energy −17.4% in the model.	Containment, leakage, and full-scale economics remain unresolved.
[2] Transient spray	Physics-based droplet model + substrate finite differences	Tracks film, transition, and nucleate boiling; smaller droplets and higher mass flow generally speed cooling.	Needs local spray statistics; drop-interaction correction has limited direct validation.
[3] Spray quenching	Iron-disk experiment at 0.1–0.5 MPa	Quench temperature rises with pressure; film-boiling heat rate rises about 105% from 0.1 to 0.5 MPa.	Polished, oxidation-limited test surface; industrial surfaces vary.
[4] Cooling review	Review of air, water, equipment, and intelligent management	Frames cooling as a coordinated system of air distribution, water temperatures, components, and controls.	Examples use different sites and methods; not one controlled comparison.
[5] DASE	Physics-based cross-system simulation and optimization	Over 5% total-energy saving in summer weeks; over 46% cooling-energy saving in winter weeks.	Modeled results; seasonal denominators and site conditions differ.
[6] CSA-MPC	Five-zone case study, field data, prediction + PSO control	Inter-zone spread 7.05 to 1.02 °C; cooling supply reduced 8.73% in the case study.	Transfer requires site recalibration and operational fault handling.
[7] RDHx	CFD rack module + 30 MW annual energy model	Rack inlet about 24.1–24.3 °C; modeled cooling PUE average 1.273.	No completed field validation; paper's module size conflicts in conclusion.
[8] Nanofluid plate	CFD/FE server model + lifecycle cost estimate	Al_2O_3 improves chip temperature; water at Re = 600 is lowest cost under assumptions.	Small modeled system; abstract and conclusion disagree on minimum cost.

4.2 Quantitative findings without false equivalence

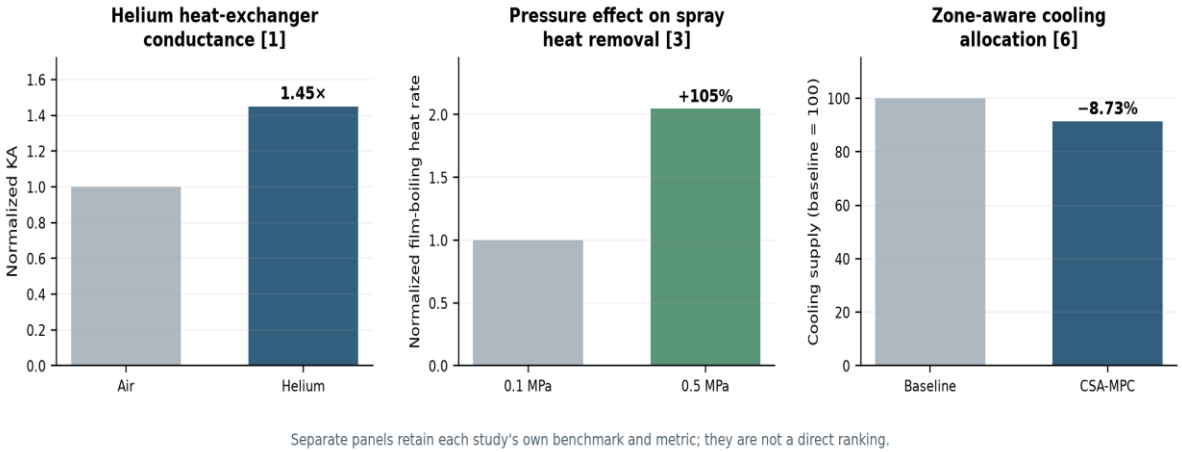


Fig. 2. Selected quantitative results from three studies. Each panel preserves the source study's own comparison and unit basis; the panels are not a common performance scale.

The chart separates three findings that are often conflated: heat-exchanger conductance, spray heat transfer, and cooling-supply reduction. Helium's 1.45× KA result is a component-level conductance comparison; the spray study's 105% change concerns film-boiling heat removal over a pressure range; and the CSA-MPC result is a cooling-supply reduction relative to one room's baseline. None is a direct estimate of the percentage reduction in whole-data-center electricity. The DASE paper's seasonal savings and the RDHx paper's cooling PUE are reported separately below because their denominators and baselines differ.

Metric	Reported result	Interpretation
Helium heat-exchanger conductance [1]	303.5 vs 209.0 W/K; 1.45×	Measured KA for the tested exchanger; not a whole-facility energy ratio.
Spray film-boiling heat rate [3]	About +105% from 0.1 to 0.5 MPa	Average rate for the reported flux and controlled iron-disk experiment.
DASE optimization [5]	>5% total energy in summer weeks; >46% cooling energy in winter weeks	Different seasons and energy denominators; retain each as reported.
CSA-MPC cooling supply [6]	8.73% lower than baseline	Reported for one five-zone case study; not equivalent to whole-site savings.
RDHx annual cooling PUE [7]	1.25–1.33; average 1.273	Modeled 30 MW reference center; not a measured fleet average.
Nanofluid economics [8]	Water at Re = 600 lowest cost; $706.98 in results/conclusion	Abstract states $706.15; source discrepancy noted.

4.3 Main design trade-offs

First, heat-transfer enhancement and energy efficiency are not identical. Helium raises measured conductance but its lower volumetric heat capacity changes the heat-carrying balance. Nanoparticles lower simulated chip temperature but raise viscosity and pumping demand. RDHx moves heat capture closer to the rack but assigns a substantial share of modeled cooling energy to pumps. The system boundary must therefore include fans, pumps, chillers, and controls before claiming an efficiency gain [1,7,8].

Second, operating conditions are decisive. Free cooling depends on outdoor temperature and humidity; chilled-water setpoints affect chiller efficiency but must remain compatible with IT inlet conditions; and MPC benefits from accurate zone-level sensing and models [4–6]. A technology that performs well under one climate, workload, or rack density may not be optimal under another. Climate-aware controls and measured operational baselines are essential for credible comparisons.

Third, spray cooling evidence is mechanistic and promising but indirect for data centers. Transient models and controlled quenching experiments improve understanding of droplet impacts, boiling regimes, and pressure effects [2,3]. Before these approaches can be recommended for IT equipment, direct tests must address electrical isolation, coolant compatibility, corrosion, contamination, spray uniformity, leak containment, serviceability, and long-duration reliability. The present set does not establish spray cooling as a ready-to-deploy data-center solution.

5. FUTURE SCOPE

Future work should connect component-scale experiments to full system energy accounting. For helium, this means multi-rack demonstrations with measured leakage rates, gas recovery, service procedures, and lifecycle cost. For spray cooling, experiments should report local droplet size, velocity, and number-flux distributions alongside surface condition and transient temperature, then test representative electronics packages and dielectric or otherwise compatible working fluids. For nanofluid-assisted plates, long-term dispersion stability, filtration, corrosion, and pump performance should be measured under representative server duty cycles.

At facility scale, field trials should compare DASE, CRAC control, and RDHx against documented baselines across seasons while reporting IT load, total facility energy, cooling energy, water use, and thermal compliance separately. Multi-zone controllers should include uncertainty, sensor-fault handling, and safe fallback modes. A common reporting protocol would make future results easier to compare without erasing the differences between component tests, rack demonstrations, and facility-level studies.

6. CONCLUSION

The eight reviewed publications show that data-center thermal management is progressing along several complementary paths: improved gas-side heat transfer with helium; clearer spray-boiling models and pressure-dependent quenching physics; better airflow and chilled-water system design; weather-aware free cooling; multi-zone CRAC optimization; rack-level RDHx; and nanofluid-enhanced cooling plates within an immersion architecture. Reported benefits include higher exchanger conductance, faster spray heat removal under selected conditions, reduced modeled energy or cooling supply, and stable rack inlet temperatures. These results are not directly comparable because they use different methods, scales, and baselines.

The practical choice should follow the facility's heat density, climate, reliability requirements, retrofit constraints, water and energy costs, and maintenance capability. System-level measurement and field validation are needed before laboratory or simulation gains can be treated as deployable savings. The strongest overall conclusion is therefore not that one cooling method is universally superior, but that efficient data centers require coordinated component design, thermal control, and facility-level energy evaluation.

REFERENCES

[1] X. Wang, H. Ran, X. Chen, L. Wang, T. Ding, and Z. Li, “Helium-cooled high-heat-flux modular data center: An experimental and simulation study on heat transfer characteristics and energy efficiency,” Applied Thermal Engineering, vol. 299, art. 131254, 2026. doi:10.1016/j.applthermaleng.2026.131254.

[2] N. Banerjee, C. Tropea, and S. Seshadri, “Transient spray cooling: An analytic predictive approach,” Physical Review Fluids, vol. 10, art. 053602, 2025. doi:10.1103/PhysRevFluids.10.053602.

[3] Y. Kita, M. Nakamatsu, S. Hidaka, M. Kohno, and Y. Takata, “Quenching mechanism of spray cooling and the effect of system pressure,” International Journal of Heat and Mass Transfer, vol. 190, art. 122795, 2022. doi:10.1016/j.ijheatmasstransfer.2022.122795.

[4] F. Zhou, W. Gu, and G. Ma, “Advancements in data center cooling systems: From refrigeration to high performance cooling,” Energy & Buildings, vol. 320, art. 114634, 2024. doi:10.1016/j.enbuild.2024.114634.

[5] Y. Zhou, S. Li, Q. Li, F. Wei, D. Yang, J. Liu, and D. Yu, “Energy savings in direct air-side free cooling data centers: A cross-system modeling and optimization framework,” Energy & Buildings, vol. 308, art. 114003, 2024. doi:10.1016/j.enbuild.2024.114003.

[6] J. Zhao, D. Xu, X. Yuan, and D. Liu, “Cooling supply allocation model predictive control strategy of data center computer room air conditioning systems considering multi-zone thermal coupling effects,” Applied Thermal Engineering, vol. 277, art. 127103, 2025. doi:10.1016/j.applthermaleng.2025.127103.

[7] J. Cho, “Numerical analysis of rack-based data center cooling with rear door heat exchanger (RDHx): Interrelationship between thermal performance and energy efficiency,” Case Studies in Thermal Engineering, vol. 63, art. 105247, 2024. doi:10.1016/j.csite.2024.105247.

[8] F. S. Alkasmoul, A. M. Almogbel, M. W. Shahzad, and A. J. Al-damook, “Thermal performance and optimum concentration of different nanofluids in immersion cooling in data center servers,” Results in Engineering, vol. 25, art. 103699, 2025. doi:10.1016/j.rineng.2024.103699.