2022

# Cooling Technologies in Datacom Facilities: An Overview and Perspectives

Amin Isazadeh

Davide Ziviani

David E. Claridge

# Cooling Technologies in Datacom Facilities: An Overview and Perspectives

**Amin Isazadeh**[1*], **Davide Ziviani**[2], **David E. Claridge**[1]
[1] Energy Systems Laboratory, Texas A&M Engineering Experiment Station and J. Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, TX 77843, USA
[2] Ray W. Herrick Laboratories, School of Mechanical Engineering, Purdue University, West Lafayette, 47907-2099, USA
[*] amin.isazadeh@tamu.edu; dziviani@purdue.edu; dclaridge@tamu.edu

## ABSTRACT

The demand for data center and network services has been rising rapidly over the last decade. However, the power demand has become stable in recent years, owing to more efficient electronic hardware, migrating to hyperscale and cloud data centers, and more efficient cooling infrastructure, among others. This paper provides a critical overview of cooling technologies and a discussion of research gaps. Cooling technologies in datacom facilities can be broadly categorized into air-cooled and liquid-cooled systems. Overhead/underfloor air delivery, hot/cold aisle layout, and hot/cold aisle containment are the primary strategies used to optimize air cooled system performance. The raised floor architecture has been widely adopted in datacom facilities, but has substantial air flow leakage (about 25–50%). It was found that the optimal ventilation system is a hard floor design with overhead cold air delivery and hot air return duct instead of room-based supply and return. Cold-aisle containment can better reduce the maximum inlet temperature of the racks and suppress temperature rise during cooling system failure, while hot-aisle containment can provide lower average inlet temperature of the racks with smaller standard deviation and is less affected by air tightness around the servers. As rack power density rises above 10 kW/rack and heat flux beyond 100 kW/cm$^2$, conventional air-cooled systems are not a viable solution for thermal management. Liquid cooling methods like spray cooling, impingement jet, immersion cooling, liquid-cooled micro-channels, and heat pipes are among the emerging technologies to overcome the capacity limitations of air-cooled systems. Pertaining to immersion cooling, transitioning into sub-cooled two-phase flow boiling, enhancing heat transfer by adding micro structures or irregularities to create more nucleation sites and higher heat transfer surface area, and utilizing nanofluids are prominent enhancement strategies gaining attention among scholars. Submerging a power electronics module in a fluid can lead to a thermal resistance of 25% that of an air-cooled system, or 30-50% that of a liquid-cooled system like microchannel or spray cooling. Depending on the existing cooling system, overall heat load, and hot spots, the heat pipe system can serve the data center as a stand-alone unit or in conjunction with an air-cooled system, a so-called hybrid system. Compared to typical air-cooled systems, the hybrid system can lower annual cooling load factor and energy consumption by 37-58% and 20-70%, respectively.

## 1. INTRODUCTION

Over the last decade, demand for data center and network services has risen dramatically. Growing demand for streaming content, evolving digital technologies, emerging blockchain IT infrastructure, and accelerated internet of things (IoT) growth as a result of expanded machine-to-machine (M2M) technology are several of the primary factors driving the need for next-generation data centers. Total internet users are expected to expand by 13% at a compound annual growth rate (CAGR) of 6%, from 4.7 billion in 2021 to 5.3 billion in 2023, 59% and 66% of the world population, respectively (CISCO, 2020). Online video streaming and gaming are estimated to contribute 2.9 ZB and 180 EB in traffic (Cisco Systems, 2019), respectively and are projected to account for 87% of the total internet traffic in 2022 (George Kamiya, 2020). In addition, machine-to-machine (M2M) technology is expected increase total connections from 8.9 billion in 2018 to 14.7 billion in 2023 (CISCO, 2020).

The United States has maintained its position as the world leader in the data center business, with more than one-third of all the world's data centers. The power consumed by data centers accounts for 2-5% of overall electricity usage in the United States. As a result of more energy-efficient electronic technology, migration to hyperscale and cloud data centers, and more efficient cooling infrastructure, among other factors, the demand for electricity has remained stable in recent years (Kamiya, 2020; Shehabi et al., 2016). Despite the fact that data center power usage has remained steady in recent years, the industry is responsible for a substantial share of the country's total electricity consumption (Shehabi et al., 2016). It has been reported that the thermal management and conditioning of data centers accounts for around 30-50% of their total energy consumption (Shah et al., 2017). Growing concern about rising energy prices and decreasing natural resources has resulted in a surge in interest in environmentally friendly and energy-efficient cooling innovations (Chandrasekaran and Bhavnani, 2017). This paper provides a critical overview of cooling technologies and a discussion of research gaps. To assure reliable electronic equipment operation while utilizing the least amount of power, a properly designed thermal management solution should balance component temperatures, humidity, acoustics, and equipment performance. Generally, cooling strategies in datacom facilities may be divided into two categories: air-cooled systems and liquid-cooled systems, both of which can be implemented at a variety of sizes (from chip level to room level).

Over the last decade, many scholars reviewed thermal management strategies in data centers including passive and active systems, airflow management (Ni & Bai, 2017; Nadjahi et al., 2018), performance evaluation metrics of data centers (Gong et al., 2020; Reddy et al., 2017), waste heat recovery and renewable energy integration into datacom facilities (Ebrahimi et al., 2014; Huang et al., 2020). However, a deeper understanding of global trends of datacom-related energy consumption including water and electricity usage, rack power density, cloud and hyperscale shift, performance metrics (e.g., PUE) is still lacking in the literature. Moreover, a broader overview of different airflow management techniques in air-cooled data centers and emerging liquid-cooled systems such as hybrid heat pipe and immersion cooling systems is also needed. This paper aims at covering these aspects in a comprehensive way and provides the reader with perspectives on research gaps.

## 2. GLOBAL TRENDS IN DATA CENTERS

Energy consumption in a data center is basically attributed to power infrastructure including power subsystems and uninterrupted power supply (UPS), cooling systems, and peripheral systems like lighting, and IT infrastructure including servers, storages, and networking. The IT infrastructure is typically the dominant energy user in any datacom facilities as shown in Figure 1 (Shehabi et al., 2016). Water consumption of data centers has recently been gaining attention. Water usage in the process of electricity generation at a primary power plant and water consumption for the internal cooling system are the main drivers of the water demand (Patterson et al., 2011). By factoring the water losses at both hydroelectric and thermoelectric power plants, a national weighted average of 7.6 liters (2.0 gallons) of fresh water is evaporated per kWh consumption by end consumers (Torcellini et al., 2003). The water consumption of US data centers is shown in Figure 2 (Shehabi et al., 2016).

Power Usage Effectiveness (PUE) and Water Usage Effectiveness (WUE) are the two primary and commonly used metrics of overall facility performance developed by the Green Grid corporation (ASHRAE, 2019); Shehabi et al., 2016). PUE measures how effectively a datacom facility delivers power to its infrastructure, whereas WUE indicates the facility's reliance on water resources (ASHRAE, 2019). PUE is defined as the ratio of overall facility power consumption to IT infrastructure power consumption and ranges between 1.0 and infinite with an optimal value of 1.0. The power utilization efficiency (PUE) of a facility is the ratio of overall facility power consumption to IT infrastructure power consumption (Voort et al., 2017). PUE was developed to collect data on the effectiveness of changes made to a data center's infrastructure. According to the United States International Trade Commission, there are approximately 8,000 data centers in the world as of January 2021, with the United States, the United Kingdom, Germany, and China accounting for approximately 33%, 6%, 5%, and 5% of total data center capacity, respectively (Daigle, 2021). According to the results of an Uptime Institute study of 846 end customers across the world, data centers have been somewhat less efficient in recent years in terms of average power use effectiveness; 1.58 in 2018, 1.67 in 2019, and 1.59 in 2020, respectively. In Figure 3, a significant improvement in efficiency was observed between 2007 and 2013, primarily achieved through cost-effective measures such as aisle containment, while improvements thereafter became minimal due to the costly and complicated new energy-saving measures implemented after that (Ascierto and Lawrence, 2020). Figure 4 shows that the vast majority of data centers in operation throughout the world have power utilization efficiencies (PUE) ranging between 2.00 and 2.49 (Supermicro, 2021).

## 3. AIR-COOLED SYSTEMS

Air-cooled systems dissipate the heat from the electronics by delivering cold air to the equipment and then exhaust hot air. Conventional computer-room air handling units (CRAH) and computer-room air conditioners (CRAC) are two commonly used systems in a datacom facilities, where CRAC units are equipped with dedicated compressors unlike the CRAH unit, which only passes air over a cooling coil supplied by a separate on-site chiller. Within a datacom facility, the two most common techniques of airflow distribution are underfloor air supply and overhead air delivery, respectively. Telecommunications rooms have traditionally lacked a raised floor and relied on overhead ducted air delivery, whereas datacom facilities have raised-floored systems that serve as supply air plenums. In the case of underfloor air delivery, perforated tiles with varying grate concentrations and layouts are utilized. Standard tiles with 22-35% airflow and high output tiles with 55-66% air grates are available (Pro Access Floors Co., 2021). The tiles may also be equipped with blowers and/or dampers to provide superior control of the supplied airflow rate. Hard floor design is a common architecture in data centers with centralized air handling units instead of CRAH/CRAC units mostly used in raised-floor configuration (ASHRAE, 2019).

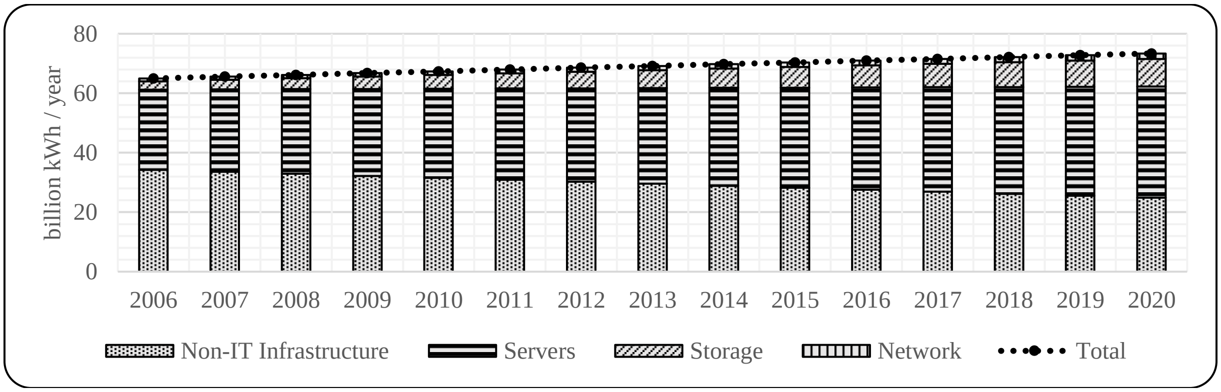


**Figure 1:** Annual electricity consumption in US data centers

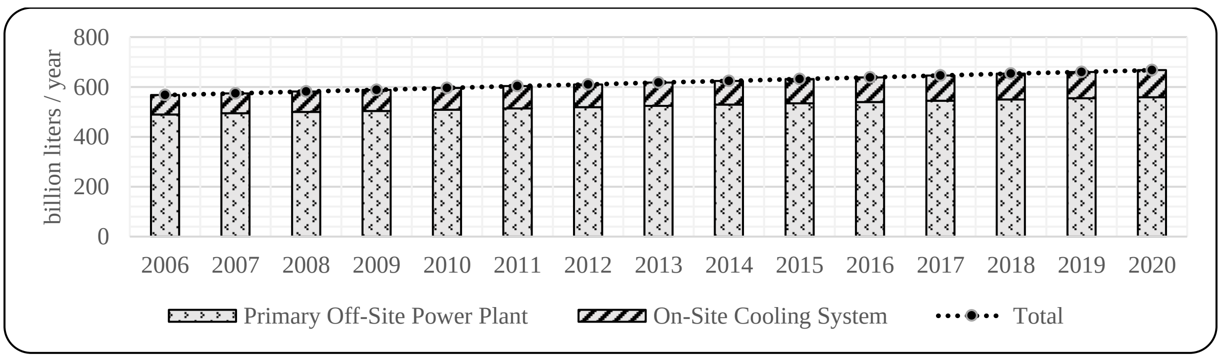


**Figure 2:** Annual water consumption in US data centers

Although the raised floor design has been widely adopted throughout the datacom industry due to its ability to provide relatively uniform air distribution along the rack, it has historically had a significant air leakage issue (ASHRAE, 2019) with about 25–50% air leakage in a traditional raised floor with uncontained aisles (Niemann et al., 2013). The primary adverse effect of air leakage is the mixing of cold and hot air, which reduces the cooling system's efficiency. Air mixing can occur by hot air recirculation and/or cold air bypass (ASHRAE, 2019; Xiong et al., 2019; Bob Sullivan et al., 2019). Hot air recirculation occurs when hot air leaving the servers recirculates to the server's inlet, where it mixes with the cold air flowing in and raises the intake temperature, resulting in higher load on the cooling system. Besides hot air recirculation, cold air bypass may occur when supplied cold air bypasses the hardware and combines with hot air leaving the equipment, lowering the temperature of the return air. By decreasing the system temperature difference, the lowered return air temperature reduces the cooling capabilities of the cooling coils, necessitating an increase in airflow to meet the load (ASHRAE, 2019).

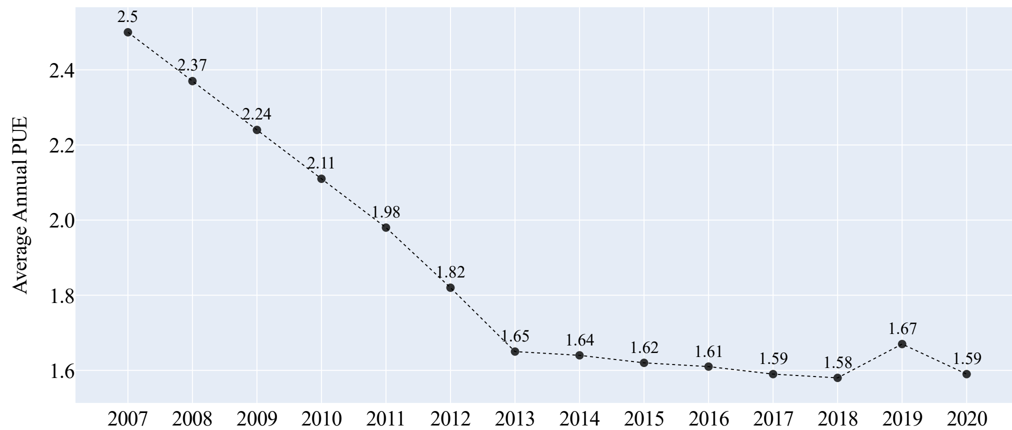


**Figure 3:** Global average annual power usage effectiveness (PUE) of data centers (Ascierto & Lawrence, 2020)

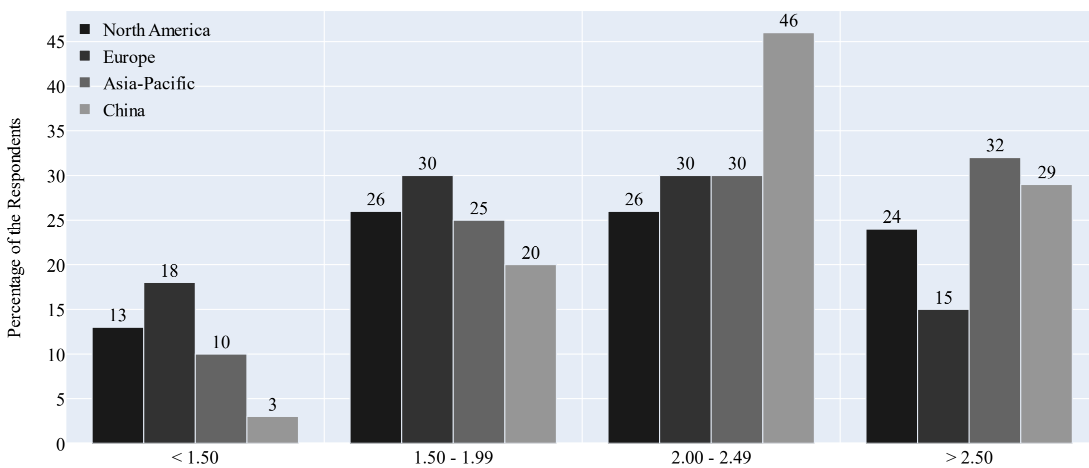


**Figure 4:** Power usage effectiveness (PUE) distribution across the globe (Supermicro, 2021)

For the cooling system to operate more efficiently, it is necessary to separate the supply and return air flows. Additionally, it is necessary to isolate the hardware's cold inlet air from the hot outgoing air to reduce the aforementioned air mixing phenomena and provide more energy efficient and effective air-cooling. The hot/cold aisle layout and confinement are two well-known tactics for separating air flows. Data center air containment systems are the most effective information technology cooling solutions (Niemann et al., 2013) and have emerged as a popular method of improving cooling system efficiency while simultaneously reducing energy consumption (ASHRAE, 2019). The containment strategy, whether cold aisle or hot aisle, can improve cooling system efficiency as well as electronics performance and reliability by increasing the predictability and efficiency of cold air distribution (Cho, 2021), providing more uniform hardware inlet air temperature, and preventing local hot spots that are common in uncontained aisles (Lin et al., 2013). In general, containment solutions improve the thermal performance of a facility in the following ways (Niemann et al., 2013; Nishimura et al., 2019; Tradat et al., 2018):

- **Higher supply air temperature** – in uncontained designs, supply air temperatures are kept lower (about 13ºC) than in contained designs (18ºC) due to the presence of frequent local hot spots created by hot air recirculation and cold air bypass in the uncontained configuration, as opposed to contained designs.
- **Increased economizer hours** – as a result of higher supply air temperature in the contained design, the cooling system may work in economizer mode for longer periods of time.
- **Prevention/minimization of local hot spots** –the containment design effectively matches the supply air temperature and the hardware inlet air temperatures, thereby eliminating the possibility of hot spots resulting from nonuniform temperature distribution along the racks.
- **Lower risk of condensation** – containment permits the supply air temperature to be well above the dew point to avoid any condensation problems, which reduces the need for dehumidifiers and humidifiers.

- **Lower cooling system capacity** – because of the possibility of raising the supply air temperature, it may be possible to downsize the cooling system, which may result in a more efficient system with reduced losses.
- **Less air leakage** – there is normally 3-10% air leakage in a contained design, but 25-50% air leakage is common in an uncontained architecture.
- **Lower air flow rate** – the heat load requires a lower air flow rate in a contained design due to better match between supply air temperature and server intake air temperature, as compared to an uncontained design.
- **More stable cold aisle temperature** – during a power outage, the inlet air temperature of the hardware rises significantly faster in an uncontained design than in a closed design.

Numerous studies have addressed the effectiveness of a containment design. Arghode et al. (2013) studied thermal characteristics of open and cold-aisle containment under over-provisioned (higher supplied airflow than the airflow within the racks) and under-provisioned (UP, lower supplied airflow than the airflow within the racks) cases by comparing computational fluid dynamics results of the open aisle design and thermal field measurements of the contained aisle design. They found that the thermal performance of cold-aisle containment is significantly superior to open aisle design by improving cold air delivery, matching the tile and rack air flow rates, more uniform temperature distribution in the cold aisles and at the server inlets, and lower air confinement at the aisle entrance caused by high speed leaving air from perforated tiles. Sundaralingam et al. (2015) experimentally studied different cold aisle containments, open aisle, partially contained, and fully contained cold aisles. They suggested that an over-provisioned fully contained design is the most effective architecture, but top only partial contained design regardless of the supplied airflow is the best design choice in the existence of geometrical or cost limitations. Ham and Jeong (2016) experimentally and numerically compared thermal management performance and air leakage of contained cold aisle with uncontained design in a modular data center when using an integrated water-side economizer. Their results showed that containment benefits data centers in terms of thermal management, air leakage, and energy savings. Tradat et al. (2018) experimentally studied ride through time (RTT) of servers installed in the contained cold aisle and analyzed the necessity of a pressure relief mechanism (by opening/closing the containment doors) during blower failure. They concluded that pressure relief is not required regardless of server type. In contrast, when the blowers failed, cold-aisle containment (CAC) kept the servers cold longer since the confinement created negative pressure between the hot stream and cold stream, allowing the servers to draw cold air from the underfloor plenum and increasing the RTT. Nishimura et al. (2019) experimentally evaluated thermal performance of partial cold-aisle containment (some but not all cold aisles are contained) by monitoring the cold aisle temperatures. They concluded that such partial containment architecture should be avoided because the uncontained cold aisles featured faster server intake temperature rise (reaching 40ºC in 8 minutes versus 34 minutes in the contained aisles), frequent hot air recirculation, and being highly influenced by higher server heat dissipation compared to the contained cold aisles. Cho (2021) experimentally and numerically studied thermal performance of both types of containment and noted that CAC requires greater design work since each aisle width must be tailored to the projected load while HAC does not have this difficulty, and its simplicity of design makes it the preferred choice. Their results showed that, contrary to the previous studies' observations, CAC led to a higher average inlet temperature with a higher standard deviation among all racks, resulting in lower cooling efficiency compared to HAC.

## 4. LIQUID-COOLED SYSTEMS

Liquid cooling is the process of removing heat (i.e., cooling) through the use of liquid, rather than air (ASHRAE, 2019). Traditional air cooling solutions will be inefficient for thermal management of high-power density packages, since heat flux for high-performance chips exceeds 100 W/cm$^2$ at the 14 nm manufacturing node (Chandrasekaran and Bhavnani, 2017) and is anticipated to reach 1000 W/cm$^2$ in the future (Nakayama, 2014). A larger air-cooled system is required to handle this problem, which will result in higher operating and maintenance costs, as well as higher energy expenditures, and a larger overall facility area. Instead, liquid-cooled technologies with greater heat removal capability might be employed (Kanbur et al., 2020). Liquid cooling is more ecologically friendly, scalable, and more targeted than air cooling (Kriech, 2020). Because of its better thermal qualities, liquid cooling may be utilized to cool the electronics either directly or indirectly, depending on the application (Nimkar et al., 2003; Kanbur et al., 2020). Jet impingement, spray cooling, and single/two phase immersion cooling are the primary types of direct liquid cooling, while heat pipe and thermosyphon are the two prominent indirect liquid cooling methods.

## 4.1 Immersion Cooling Systems (ICS)

Immersion cooling (IM) systems entail immersing the electronics in a non-conductive, non-flammable dielectric liquid (like mineral oil and fluoroketones) bath. The fluid, as well as the hardware, are housed in a leak-proof container (Kriech, 2020). Depending on the boiling point of the dielectric fluid, immersion cooling can be either single phase or two phases where the fluid evaporates and condenses back into the main fluid tank. Immersion cooling is more compact (Matsuoka et al., 2017) and provides more energy savings (Chandrasekaran et al., 2017) compared to conventional air-cooled data centers. Either a fully enclosed liquid tank or an open liquid bath can be used to submerge data communication equipment. As compared to alternative liquid and hybrid cooling systems, open bath immersion cooling (OBI) offers a number of advantages, including less leakage risk, lower power demand, limited natural resource usage, less complex design, lowered greenhouse gas emissions, more uniform temperature distribution, more efficient heat dissipation, and cost-effective construction (Tuma, 2010).

Beyond the fact that immersion cooling may be conducted in either an open or closed liquid bath, the liquid itself can undergo phase transition to give increased heat dissipation by utilizing its latent heat capacity, which results in higher heat dissipation. Single-phase liquid-cooled systems are less complicated to build and have a high heat transfer efficiency, making them popular (Tiwari et al., 2012). However, two-phase cooling systems provide a number of benefits over single-phase options that make them more appealing. Some of the major advantages are higher heat transfer coefficients, more uniform temperature distribution, suitable for high power electronic hardware, and being cost effective (Seuret et al., 2018; Kanbur et al., 2020; Barnes and Tuma, 2010). Over the last few decades, numerous studies have been conducted to enhance heat removal capacity of such cooling systems by incorporating strategies like subcooling, better working fluids, utilizing flow guides, micro-structure surfaces like micro-porous to create more nucleation sites and higher heat transfer surface area (Namazi et al., 2022). Some of these studies are reviewed below.

Nimkar et al. (2003, 2006) experimentally studied pool boiling of a vertically mounted silicon surface on a glass substrate with etched micro-pyramidal reentrant cavities with mouth size of 40 µm, immersed in FC-72 dielectric fluid. They achieved a critical heat flux of 12.8 W/cm$^2$, less than the expected 15.0 W/cm$^2$ estimated by Zuber correlation. It was discovered that bubble departure diameter and frequency increase as the heat flux rises. Their results also indicated that pyramidal cavities with sharp internal corners outperformed bulbous cavities with smooth inner surface. Medavaram et al. (2005) utilized a high speed camera to capture the effect of heat flux on bubble departure diameter and frequency under subcooled and saturated conditions. It was found that increasing subcooling decreases both frequency and departure diameter, but significantly raises critical heat flux. Gess et al. (2014) experimentally studied small form factor pool and flow boiling of two different surface enhanced silicon die (microporous and microfinned) immersed in Novec-649 dielectric. The microporous surface was prepared by sintering copper microporous with nominal sphere size of 10-50 µm and the microfinned surface was composed of square copper fins with 400 µm on each side and height of 2 mm. Under the pool boiling condition, it was discovered that the heat transfer coefficient of the enhanced surfaces were about three times higher than that of bare silicon with 18°C temperature drop at the measured highest heat fluxes. In the case of flow boiling, the achieved maximum heat dissipations were twice as the one measured for pool boiling. Their experimental results also indicated that the microfinned structures outperform the microporous surfaces at higher dielectric flow rates, while the microporous surface performed better in lower flow rates. Chandrasekaran et al. (2017) experimentally investigated the effects of subcooling, mass flow rate, and surface enhancement on the flow boiling of a compact cooling system using vertical arrays of heaters simulating electronic chips on a PCB immersed in Novec-649. It was found that lowering the facility water temperature reduces the surface temperature in the forced convection and partial boiling regimes (i.e., the fully developed boiling regime is not affected) while raising the flow rate has no effect on the slope of the boiling curve for the bare silicon surface. The boiling curves of microporous surfaces exhibited a similar behavior to variations in flow rate and facility water temperature as those of bare silicon. However, for the microfinned surface, increases in flow rate and decreases in facility water temperature both reduced surface temperature in the forced convection and partial boiling regimes with no impact on the fully developed boiling regime. By maximizing the flow rate and minimizing the facility water temperature, heat fluxes of 20.5 W/cm$^2$, 20.8 W/cm$^2$, and 22.8 W/cm$^2$ were achieved for the bare silicon, microfinned surface, and microporous surface, respectively. It was also discovered that the enhanced surfaces led to significantly lower surface temperatures compared to the bare silicon surface. The authors concluded that the microporous heat sink is the preferable choice in bringing higher heat dissipation with lower surface temperature. Matsuoka et al. (2017) numerically and experimentally studied different dielectric fluids and natural convection in a liquid bath thermally supported by four cooling plates (three parallel to the side walls and one on the bottom of the tank) and populated by six server units. Silicon oil 20 and 50, soybean oil, FC-43, and FC-3283 are the fluids chosen for this study. It was discovered that heat dissipation capability of a system with natural convection is greatly impacted by the fluid viscosity such that less viscous liquid performs better. Among the studied fluids, the Fluorinerts (FC-43 and FC-3283) outperformed the viscous oils by providing higher heat flux dissipation and lower fluid temperature. It was also numerically found that the CPUs' surface temperatures fall and the flow rate over the CPUs' surfaces rises as the Rayleigh number increases. The effect of changing the characteristics of a certain section, including the CPU workload and slot removal, is restricted to that portion alone; the impact on other locations is negligible. These features are particularly good in terms of the data center's operational stability. It was experimentally explored that the CPUs' junction temperatures linearly increase with the CPU power and a finned heat sink attachment to the CPUs' surfaces can double up the amount of heat dissipation. Eiland et al. (2014) examined flow boiling immersion cooling with the use of mineral oil and found the partial power usage effectiveness of 1.03 to 1.17 depending on the operating temperature of the serves. Alsaati et al. (2020) experimentally identified and characterized boiling regimes for confined boiling in a passive immersion cooling and examined heat source size on the confined zone. It was found that heat dissipation occurred by conduction and natural convection at low heat fluxes like unconfined boiling. The average surface temperature was relatively lower in confined boiling compared to unconfined boiling. The optimal thermal performance was observed with a confinement gap size corresponding to 0.2 Bond number. Despite a significant drop in surface temperature, the CHF for this optimal confinement gap size was only 27% of the unconfined CHF. When the confined region was extended to an area bigger than the heater, even greater decreases in CHF were found.

## 4.2 Heat Pipe (HP) Cooling Systems

A heat pipe is a closed conduit that transports a working fluid that evaporates at the tunnel's end (evaporator) by heat absorption and condenses at the other end of the tunnel (condenser") due to heat release. As a terminal unit, heat pipes have been utilized to increase heat dissipation rates by providing excellent heat transfer performance, good geometrical compatibility, low risk of coolant leakage, and varied scaling applications at the rack, server, and chip levels. Heat pipes are mainly classified into standard heat pipes (HP) with constant or variable conductance with capillary as the main driving force, thermosyphons (TS) whose driving force is gravity, and heat pipes with separated liquid and vapor lines (so called loop heat pipe or LHP and loop thermosyphon or LTS depending on the driving force).

Standard heat pipes and thermosyphons with shared liquid and vapor lines have been commonly used for small-scale heat removal such as server or chip level cooling due to their limited cooling capacities, while heat pipes with separated vapor and liquid lines like LHP and LTS are typically employed in data centers at larger scales, ranging from room to rack level cooling to take advantage of natural free cooling provided by outside air along air-side or water-side economizers typically used air-cooled systems (Isazadeh et al., 2021). Abe et al. (2010) reported heat transfer enhancement by utilizing dilute aqueous solution and nanofluids in a single/two-phase narrow channel heat exchanger and standard copper flat heat pipe. The dilute solution is 1-butanol aqueous solution (5wt%) whose surface tension increases with temperature. The nanofluid is composed of silver nanoparticles with concentration of 0.25 mol/cm$^3$, which are synthesized by microwave-polyol method to provide stable and uniform nanofluid. The flat heat pipe has dimensions of 200 × 8.7 × 1.5 mm (length × width × thickness), whose condenser section is cooled by either single-phase narrow heat exchanger (53 × 43 × 3 mm) or two-phase narrow heat exchanger (69 × 42 × 12 mm). It was found that longer heat pipe had a higher thermal resistance and relatively lower maximum heat flux (45W for 200mm long HP compared to 70W for the 150mm long HP). The thermal performance of the heat pipes was increased 6 times compared to water-based heat pipes by the use of dilute aqueous solutions of high carbon alcohols. In the case of implementation in server rack, the proposed system featured lower electricity power demand, quiet operation, compact design, and conceivable use of server rack waste heat with the maximum heat removal capacity of more than 100 W, compared to air-cooled system. Chen and Chou (2014) experimentally studied impacts of filling ratio of acetone (99.87% pure) and leakage on performance of flat plate heat pipes (FPHPs) with 0.4mm deep capillary grooves. The capillary-driven heat pipe is made of aluminum 6061 sized 150 × 50 × 2.5 mm (length × width × thickness) and composed of 30mm long evaporator, 90mm long adiabatic section, and 30mm long air-cooled condenser. The heat pipe has 14 parallel rectangular channels, each sized 3.0 × 1.7 mm (width × height). The liquid charge and heat input varied between 5-50% and 5-60 W, respectively. Best thermal performance (lowest thermal resistance) was achieved at 25% filling ratio. At the optimal filling ratio, the maximum heat transfer capability, minimum thermal resistance, and maximum effective thermal conductivity were about 47 W, 0.254 ºC/W, and 3150 W/m-K, respectively. It was found that an improper vacuum and leakage result in substantial reduction in maximum thermal conductivities, 200-306 W/m-K and 164 W/m-K, respectively. The leakage reduced thermal conductivity by a factor of 19.2. Amin et al. (2020) investigated the performance of a liquid cooling system that incorporates a modified heat sink with four parallel heat pipe channels that distribute liquid flow uniformly across the surface of the stacked central processor unit (CPUs) for maximum heat transfer efficiency. Water was employed as a heat transfer agent in the heat pipes as well as a coolant for the condenser in this system. The simulation results showed that the proposed heat sink design reduced the surface temperature of the CPUs by 28.28% from 64.25 °C (147.65 °F) to approximately 46.08 °C (114.95 °F), allowing the server rack to accommodate more CPUs than an air-cooled system. Furthermore, according to the results of the field investigation, the system saved 92% cooling energy compared to conventional liquid-cooled data centers. Furthermore, server fans were turned off, and the temperature of the processor junction was drastically reduced, resulting in a 13 percent reduction in IT energy consumption overall.

The loop heat pipe (LHP) system was invented by Gerasimov et al. in 1975 (Gerasimov et al., 1975). The driving force of a loop heat pipe can be capillary force (Li et al., 2012; J. Li et al., 2010; Tang et al., 2012), gravity (Jouhara and Ezzuddin, 2013), or mechanical pump (Crepinsek and Park, 2012; Jie et al., 2008). The loop heat pipe systems are extremely reliable and need little maintenance, making them ideal for use in space and computer cooling applications. Zhang et al. (2015) experimentally examined thermal performance, mass flow rate, and flow resistance features of a pump-driven loop heat pipe system comprises five tube-fin heat exchangers as evaporators and 3 tube-fin heat exchangers as condensers with the working fluid of R-22 and filling ratio of 42.9%. It was found that heat transfer rate and coefficient of performance (COP) are both increasing substantially in response to the rise of indoor-outdoor temperature differences; when the indoor-outdoor temperature difference was 10°C, the unit's coefficient of performance was 3.75, then climbed to 9.37 when the temperature difference was 25°C. It was also observed that increasing the flow rate from 200 to 1,100 kg/h of R-22 for a single evaporator/condenser initially raises the heat transfer rate up to a certain flow rate and then the heat transfer rate reduces. At lower flow rates, the heat transfer was affected by the rise in the flow velocity, while the heat dissipation rate was suppressed by high pressure drop and lower temperature difference between the evaporator and indoor air at higher flow rates. The flow resistance and pump power both reduced dramatically when the mass flow rate declined, resulting in a higher COP. It was also discovered that the heat transfer rate was slightly influenced by the mass flow rate when the vapor quality at the evaporator is between 0.3 and 0.6. The temperature differential between the evaporator inlet and outlet, as well as the sensible heat ratio, elevated as the mass flow rate in the PLHP increased. They finally reported that at mass flow rate of 1,400 kg/h and indoor-outdoor temperature difference of 10 to 30°C, the heat transfer rate was increasing linearly with respect to the temperature difference. Zhou et al. (2018) examined energy efficiency ratio and thermal characteristics of a pump-driven loop heat pipe system installed in a 40 m2 data center in Beijing, China. The PLHP, whose operation was limited to outside air dry-bulb temperature ≤ 20°C and indoor-outdoor temperature difference ≥ 5°C, consisted of an indoor evaporator, two outdoor condensers, and a liquid magnetic pump with R-32 as the refrigerant. The results revealed that the combination of PLHP and air conditioning system can annually save energy over 30% while maintaining an indoor temperature between 18°C and 25°C with average payback of 3.9 years. It was also noted that the proposed cooling system is suitable for 74.2% of Chinese cities according to the specified operating conditions. It was also discovered that the energy efficiency ratio (EER) of the unit drops as the outside air temperature increases.

## 5. CONCLUSIONS

Air-cooled systems are the dominant cooling method in data centers. The overhead/underfloor air delivery, hot/cold aisle layout, and hot/cold aisle containment are the primary strategies used to optimize air cooled system performance by eliminating/minimizing hot air recirculation and/or cold air bypass. The raised floor architecture has been widely adopted in datacom facilities, but has substantial air flow leakage (about 25–50%). It was found that the optimal ventilation system is a hard floor design with overhead cold air delivery and hot air return duct instead of room-based supply and return. Cold-aisle containment can better reduce the maximum inlet temperature of the racks and suppress temperature rise during cooling system failure, while hot-aisle containment can provide lower average inlet temperature of the racks with smaller standard deviation and is less affected by air tightness around the servers. Aisle containment can reduce annual cooling load factor and PUE by 40-50% and 16%, respectively, compared to open design.

As rack power density rises above 10 kW/rack and heat flux beyond 100 $kW/cm^2$, conventional air-cooled systems are not a viable solution for thermal management. Liquid cooling methods like spray cooling, impingement jet, immersion cooling, liquid-cooled micro-channels, and heat pipes are among the emerging technologies to overcome the capacity limitations of air-cooled systems. Improvements in thermal management and energy performance by utilizing these technologies are summarized below:

- In immersion cooling, transitioning into sub-cooled two-phase flow boiling, enhancing heat transfer by adding micro structures or irregularities to create more nucleation sites and higher heat transfer surface area, and utilizing nanofluids are prominent enhancement strategies gaining attention among scholars. Immersion cooling can lead to junction-to-fluid resistivity of 0.15 °C/W-cm$^2$ and heat removal capacity beyond 400 W/cm$^2$, indicating that a power electronics module using this cooling technique might have a thermal resistivity of 25% of a conventional air-cooled system and 30-50% of a conventional liquid cooled technology like microchannel and spray cooling. In addition, immersion cooling with micro-structured surfaces including reentrant cavities, micro-fins, or micro-porous can raise the heat dissipation rate and heat transfer coefficient, at a given fluid superheat, by up to 4 times and 330%, respectively, compared to bare surfaces. Moreover, nanofluids can raise heat transfer coefficient by 6–11% compared to that of the base fluid. Eventually, small flow rate like 660 mL/min can increase the maximum heat dissipation by 40% and 70% with bare silicon and surface enhancement, respectively, compared to pool boiling.

- Loop heat pipe systems are widely used in data centers and are heavily climate dependent so that their operations are limited by indoor-outdoor temperature difference (good for cold climates). However, hybrid cooling systems and evaporative condensers are two common solutions in extending the operating temperature range of such systems. In hybrid cooling system, heat pipe can be integrated with air conditioner or into vapor compression cycle. The hybrid system can reduce annual cooling load factor and annual energy consumption by 37-58% and by 20-70%, respectively, depending on the climate zone, compared to conventional air-cooled systems. Furthermore, evaporative condenser can increase annual free cooling time by 7-14%, with the impact being more prominent in locations with drier weather and also can enhance the heat transfer by 7% to 33%, which is more pronounced at smaller indoor-outdoor temperature differences.

In conclusion, moving towards green data centers with PUE of less than 1.25 (Supermicro, 2021) requires incorporating energy-efficient and sustainable thermal management technologies and highly efficient IT equipment, solar power and other forms of alternative fuels, together with cutting-edge electronic components that maximize efficiency. Datacom facilities are multi-scale in nature, including multiple size scales from chip level to facility level (Samadiani, 2009). Individually tailored thermal management systems are possible for each scale. However, these scales are interrelated, and the thermal performance of one scale directly affects the performance of the others. Therefore, it becomes necessary to optimize the operating systems together to achieve the green PUE values. Figure 5 depicts various technologies that can be utilized at different scales of a datacom facility.

It should be noted that choosing the suitable cooling system is strongly dependent on the datacom facility's service time (short-term or long-term), size of the facility, climate, and geographical location. Liquid-cooled systems are superior to air-cooled systems in terms of providing higher heat removal capacity, more uniform and lower surface temperature, and lower overall energy usage, but implementation of such cooling systems may be easier at the very early stage of building the facility. Retrofits such as new duct work for hot-aisle containment or exhaust chimney, pipe lines and liquid-cooled heat sinks for liquid-cooled system can be problematic. It is worth noting that immersion cooling is an emerging and advanced cooling technology and predicted to be a viable and effective cooling solutions in the near future. The major effort in this area is engineering a cost-effective dielectric fluid with high heat removal capacity with no/minimal detrimental impact on the submerged hardware and through understanding of the pool/flow boiling mechanism within the fluid bath when multiple electronic components are placed next to each other due to potential counter-effects. On the other hand, the heat pipe system is widely adopted in data centers, but this application is very dependent on climate conditions. Recent efforts in this area combined the heat pipe system with other cooling technologies.

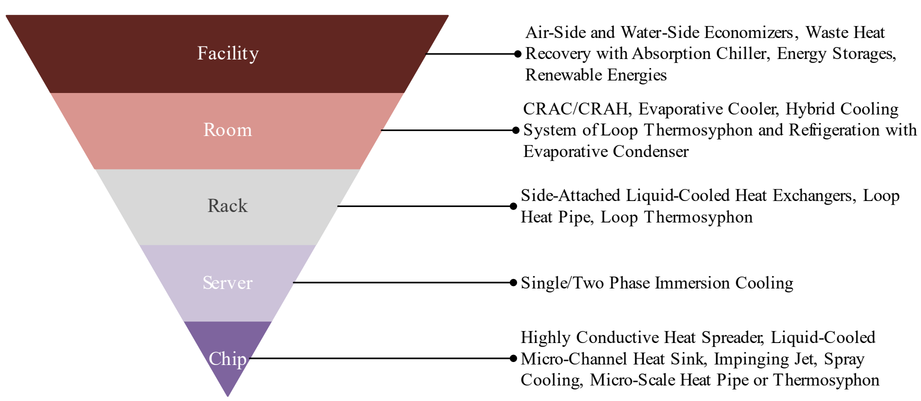


**Figure 5.** Different energy reduction technologies categorized in terms of scale of application.

# REFERENCES


Abe, Y., Fukagaya, M., Kitagawa, T., Ohta, H., Shinmoto, Y., Sato, M., & Iimura, K. (2009). Advanced Integrated Cooling Systems for Thermal Management in Data Centers. *ASME 2009 InterPACK Conference, Volume 2*, *2*, 619–624. https://doi.org/10.1115/InterPACK2009-89009

Alsaati, A. A., Weibel, J. A., & Marconnet, A. M. (2020). Confined Immersion Cooling in Microscale Gaps. *2020 19th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, *2020-July*, 475–481. https://doi.org/10.1109/ITherm45881.2020.9190616

Amin, M., Abdel Aziz, G. A., Naraghi, M., Palatty, M., Benz, A., & Ruiz, D. (2020). Improving the Data Center Servers Cooling Efficiency via Liquid Cooling-based Heat Pipes. *2020 IEEE Industry Applications Society Annual Meeting*, 1–5. https://doi.org/10.1109/IAS44978.2020.9334731

Arghode, V. K., Sundaralingam, V., Joshi, Y., & Phelps, W. (2013). Thermal Characteristics of Open and Contained Data Center Cold Aisle. *Journal of Heat Transfer*, *135*(6), 1–11. https://doi.org/10.1115/1.4023597

Ascierto, R., & Lawrence, A. (2020). *Data Center 2020 Survey*. https://doi.org/UII-38 v1.1

ASHRAE. (2019). *HVAC Applications - Chapter 20: Data Centers and Telecommunication Facilities* (SI). ASHRAE.

Bob Sullivan, R., Li, G., & Zhang, X. (2019). Cold aisle or hot aisle containment - Is one better than the other? *INTELEC, International Telecommunications Energy Conference (Proceedings)*, *2018-Octob*. https://doi.org/10.1109/INTLEC.2018.8612444

Chandrasekaran, S., & Bhavnani, S. (2017). Flow distribution and nucleation suppression in a small form factor liquid immersion cooled server model. *ASME 2017 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems*. https://doi.org/10.1115/IPACK2017-74025

Chandrasekaran, S., Gess, J., & Bhavnani, S. (2017). Effect of subcooling, flow rate and surface characteristics on flow boiling performance of high performance liquid cooled immersion server model. *Proceedings of the 16th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, ITherm 2017*, 905–912. https://doi.org/10.1109/ITHERM.2017.7992582

Chen, J.-S., & Chou, J.-H. (2014). Cooling performance of flat plate heat pipes with different liquid filling ratios. *International Journal of Heat and Mass Transfer*, *77*, 874–882. https://doi.org/10.1016/j.ijheatmasstransfer.2014.06.029

Cho, J. (2021). Hot-aisle and cold-aisle containemnt comparing thermal performance for data center cooling. *ASHRAE Journal*, *August*, 11.

CISCO. (2020). *Annual Internet Report (2018-2023)*. https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

Cisco Systems, I. (2019). Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper. *Cisco Forecast and Methodology*, 2017–2022. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html

Crepinsek, M., & Park, C. (2012). Experimental analysis of pump-assisted and capillary-driven dual-evaporators two-phase cooling loop. *Applied Thermal Engineering*, *38*, 133–142. https://doi.org/10.1016/j.applthermaleng.2012.01.022

Daigle, B. (2021). *Data Centers Around the World: A Quick Look* (Issue May).

Ebrahimi, K., Jones, G. F., & Fleischer, A. S. (2014). A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat recovery opportunities. *Renewable and Sustainable Energy Reviews*, *31*, 622–638. https://doi.org/10.1016/j.rser.2013.12.007

Eiland, R., Fernandes, J., Vallejo, M., Agonafer, D., & Mulay, V. (2014). Flow Rate and inlet temperature considerations for direct immersion of a single server in mineral oil. *Fourteenth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, 706–714. https://doi.org/10.1109/ITHERM.2014.6892350

George Kamiya. (2020). *Data Centres and Data Transmission Networks*. International Energy Agency (IEA_. https://www.iea.org/reports/data-centres-and-data-transmission-networks#%0Ahttps://www.iea.org/reports/data-centres-and-data-transmission-networks

Gerasimov, Y. F., Maidanik, Y. F., Shchegolev, G. T., Filippov, G. A., Starikov, L. G., Kiseev, V. M., & Dolgirev, Y. E. (1975). Low-temperature heat pipes with separate channels for vapor and liquid. *Journal of Engineering Physics*, *28*(6), 683–685. https://doi.org/10.1007/BF00867371

Gess, J., Bhavnani, S., Ramakrishnan, B., Johnson, R. W., Harris, D., Knight, R. W., Hamilton, M., & Ellis, C. (2014). Impact of surface enhancements upon boiling heat transfer in a liquid immersion cooled high performance small form factor server model. *Thermomechanical Phenomena in Electronic Systems -Proceedings of the Intersociety Conference*, 435–443. https://doi.org/10.1109/ITHERM.2014.6892314

Gong, X., Zhang, Z., Gan, S., Niu, B., Yang, L., Xu, H., & Gao, M. (2020). A review on evaluation metrics of thermal performance in data centers. *Building and Environment*, *177*(April), 106907. https://doi.org/10.1016/j.buildenv.2020.106907

Ham, S. W., & Jeong, J. W. (2016). Impact of aisle containment on energy performance of a data center when using an integrated water-side economizer. *Applied Thermal Engineering*, *105*, 372–384. https://doi.org/10.1016/j.applthermaleng.2015.05.069

Huang, P., Copertaro, B., Zhang, X., Shen, J., Lofgren, I., Ronnelid, M., Fahlen, J., Andersson, D., & Svanfeldt, M. (2020). A review of data centers as prosumers in district energy systems: Renewable energy integration and waste heat reuse for district heating. *Applied Energy*, *258*. https://doi.org/10.1016/j.apenergy.2019.114109

Isazadeh, A., Kamal, R., Yagua, C., Eluvathingal, S., & Claridge, D. E. (2021). Detecting deficiencies using building performance data in healthcare facilities: Improving operational efficiency with Continuous Commissioning ®. *Energy and Buildings*, *241*, 110953. https://doi.org/10.1016/j.enbuild.2021.110953

Jouhara, H., & Ezzuddin, H. (2013). Thermal performance characteristics of a wraparound loop heat pipe (WLHP) charged with R134A. *Energy*, *61*, 128–138. https://doi.org/10.1016/j.energy.2012.10.016

Kanbur, B. B., Wu, C., Fan, S., Tong, W., & Duan, F. (2020). Two-phase liquid-immersion data center cooling system: Experimental performance and thermoeconomic analysis. *International Journal of Refrigeration*, *118*, 290–301. https://doi.org/10.1016/j.ijrefrig.2020.05.026

Kriech, M. (2020). *What you should know about data center cooling technologies*. DataCenters. https://www.datacenters.com/news/data-center-cooling-future-of-cooling-systems-methods-and-technologies

Li, H., Liu, Z., Chen, B., Liu, W., Li, C., & Yang, J. (2012). Development of biporous wicks for flat-plate loop heat pipe. *Experimental Thermal and Fluid Science*, *37*, 91–97. https://doi.org/10.1016/j.expthermflusci.2011.10.007

Lin, P., Avelar, V., & Niemann, J. (2013). Implementing Hot and Cold Air Containment in Existing Data Centers. *Data Center Science Center*.

Matsuoka, M., Matsuda, K., & Kubo, H. (2017). Liquid immersion cooling technology with natural convection in data center. *2017 IEEE 6th International Conference on Cloud Networking (CloudNet)*, 1–7. https://doi.org/10.1109/CloudNet.2017.8071539

Medavaram, G. L., Bhavnani, S. H., & Jaeger, R. C. (2005). Ebullition characteristics of an isolated surface microstructure for immersion cooled heat sinks. *IEEE Transactions on Components and Packaging Technologies*, *28*(2), 263–271. https://doi.org/10.1109/TCAPT.2005.848535

Nakayama, W. (2014). Heat in Computers: Applied heat transfer in information technology. *Journal of Heat Transfer*, *136*(1). https://doi.org/10.1115/1.4025377

Namazi, M., Nayebi, M., Isazadeh, A., Modarresi, A., Marzbali, I. G., & Hosseinalipour, S. M. (2022). Experimental and numerical study of catalytic combustion and pore-scale numerical study of mass diffusion in high porosity fibrous porous media. *Energy*, *238*, 121831. https://doi.org/10.1016/j.energy.2021.121831

Ni, J., & Bai, X. (2017). A review of air conditioning energy performance in data centers. *Renewable and Sustainable Energy Reviews*, *67*, 625–640. https://doi.org/10.1016/j.rser.2016.09.050

Niemann, J., Brown, K., & Avelar, V. (2013). *Hot-Aisle vs . Cold-Aisle Containment for Data Centers*. 1–13.

Nimkar, N. D., Bhavnani, S. H., & Jaeger, R. C. (2006). Benchmark heat transfer data for microstructured surfaces for immersion-cooled microelectronics. *IEEE Transactions on Components and Packaging Technologies*, *29*(1), 89–97. https://doi.org/10.1109/TCAPT.2005.850536

Nimkar, N. D., Bhavnani, S. H., & Jaeger, R. C. (2003). Benchmark heat transfer data for micro-structured surfaces for immersion cooled microelectronics. *Advances in Electronic Packaging*, *2*, 575–585. https://doi.org/10.1115/ipack2003-35267

Nishimura, S., Mino, Y., & Tsuda, A. (2019). Thermal environment under the partial aisle containment in data center. *INTELEC, International Telecommunications Energy Conference (Proceedings)*, *2018-Octob*. https://doi.org/10.1109/INTLEC.2018.8612295

Patterson, M., Azevedo, D., Belady, C., & Pouchet, J. (2011). Water usage effectiveness (WUE): a Green Grid data center sustainability metric. *White Paper*, *35*.

Pro Access Floors Co. (2021). *Perforated Floor Tiles*. https://www.computerfloorpros.com/raised-floor-products/raised-floor-tiles/

Samadiani, E. (2009). Energy efficient thermal management of data centers via open multi-scale design [Georgia Institute of Technology]. In *ProQuest Dissertations and Theses*. https://linkresolver.tamu.edu/tamu?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:ke

Seuret, A., Iranfar, A., Zapater, M., Thome, J., & Atienza, D. (2018). Design of a Two-Phase Gravity-Driven Micro-Scale Thermosyphon Cooling System for High-Performance Computing Data Centers. *Proceedings of the 17th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, ITherm 2018*, 587–595. https://doi.org/10.1109/ITHERM.2018.8419531

Shah, J. M., Awe, O., Gebrehiwot, B., Agonafer, D., Singh, P., Kannan, N., & Kaler, M. (2017). Qualitative Study of Cumulative Corrosion Damage of Information Technology Equipment in a Data Center Utilizing Air-Side Economizer Operating in Recommended and Expanded ASHRAE Envelope. *Journal of Electronic Packaging, Transactions of the ASME*, *139*(2). https://doi.org/10.1115/1.4036363

Shehabi, A., Smith, S. J., Sartor, D. A., Brown, R. E., Herrlin, M., Koomey, J. G., Masanet, E. R., Horner, N., Azevedo, I. L., & Lintner, W. (2016). *United States Data Center Energy Usage Report*.

Sundaralingam, V., Arghode, V. K., Joshi, Y., & Phelps, W. (2015). Experimental characterization of various cold aisle containment configurations for data centers. *Journal of Electronic Packaging, Transactions of the ASME*, *137*(1), 1–8. https://doi.org/10.1115/1.4028520

Supermicro. (2021). *Data center & the environment*.

Tiwari, M. K., Zimmermann, S., Sharma, C. S., Alfieri, F., Renfer, A., Brunschwiler, T., Meijer, I., Michel, B., & Poulikakos, D. (2012). Waste heat recovery in supercomputers and 3D integrated liquid cooled electronics. *13th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, 545–551. https://doi.org/10.1109/ITHERM.2012.6231478

Torcellini, P., Long, N., & Judkoff, R. (2003). Consumptive water use for U.S. power production. *2004 Winter Meeting - Technical and Symposium Papers, American Society of Heating, Refrigerating and Air-Conditioning Engineers*, *December*, 97–101.

Tradat, M. I., Puvvadi, U. L. N., Sammakia, B. G., Ghose, K., Ibrahim, M., Calder, A., Peddle, T., Seymour, M., & Alissa, H. A. (2018). The Impact of Cold Aisle Containment Pressure Relief on IT Availability. *Proceedings of the 17th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, ITherm 2018*, 944–952. https://doi.org/10.1109/ITHERM.2018.8419557

Tuma, P. E. (2010). The merits of open bath immersion cooling of datacom equipment. *Annual IEEE Semiconductor Thermal Measurement and Management Symposium*, 123–131. https://doi.org/10.1109/STHERM.2010.5444305

Voort, T. Van De, Zavrel, V., GALDIZ, I. T., & Hensen, J. (2017). Analysis of performance metrics for data center efficiency. *REHVA Journal*, *February*.

Zhang, S., Ma, G., & Zhou, F. (2015). Experimental Study on a Pump Driven Loop-Heat Pipe for Data Center Cooling. *Journal of Energy Engineering*, *141*(4), 04014054. https://doi.org/10.1061/(asce)ey.1943-7897.0000253

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