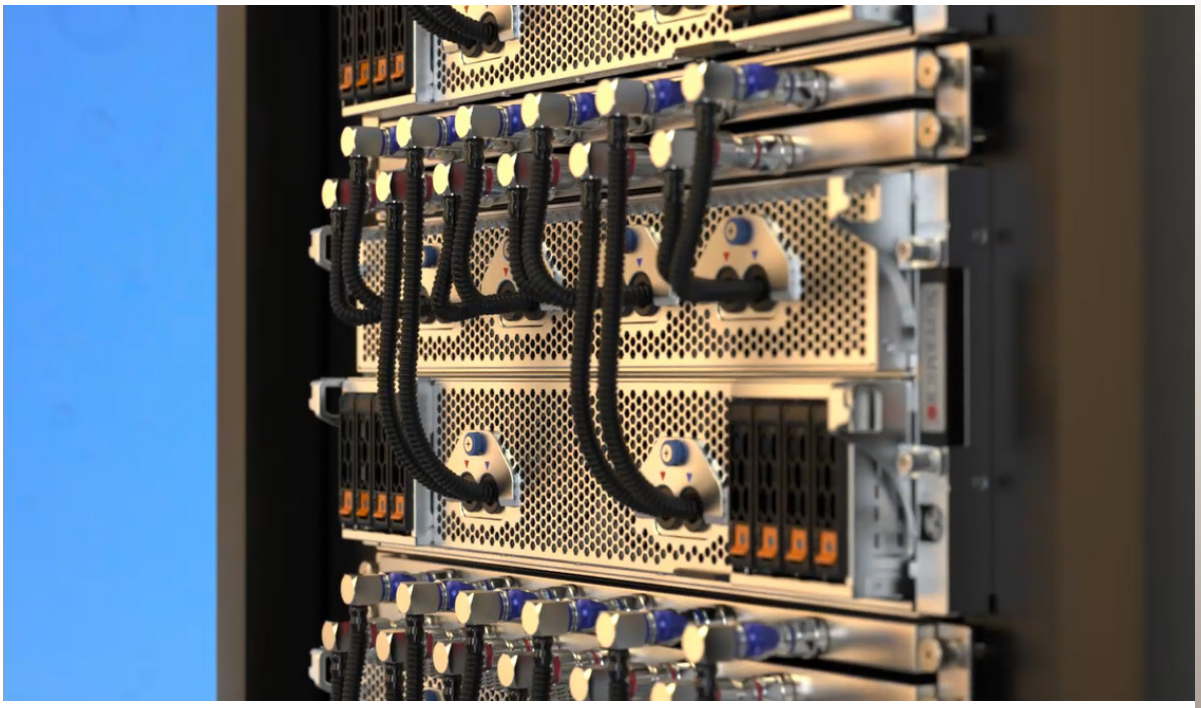
SPONSORED CONTENT | WHITE PAPER

White paper

The **top 6 reasons** you need **liquid cooling** in your **data center**



CIO

SPONSORED BY NVIDIA | SUPERMICRO

White paper

AS THE demand for high-performance computing (HPC) and artificial intelligence (AI) processing surges, data centers are being pushed to their power, efficiency, and operational cost limits.

Server racks packed with processors with high power density are presenting unprecedented operational challenges. AFCOM reported that the average rack density has increased from 8.5 kW per rack in 2023 to 12 kW per rack in 2024.[1] Densities of up to 70 kW are not unusual, and we are expecting the rack power budget to increase to 150 kW. This trend shows no signs of abating: 55% of the respondents to AFCOM's survey expect racks to become denser over the next 12 to 36 months.

The traditional method of cooling data centers — air conditioning — is no longer sufficient or efficient for the intense heat generated by the significant number of HPE and AI servers, making liquid cooling not only a better option but a requirement.

It is also inherently wasteful because air conditioning cools an entire room or data center rather than just the IT equipment inside. What's more, operators also often overinvest in air cooling to be on the safe side. The Uptime Institute has called this over-provisioning "probably a more common issue than underprovisioning due to rising rack densities."[2]

Preparing for tomorrow's cooling demands

As data center managers plan for the future, they need to consider trends in rack compute and power density and the toll air conditioning takes on power usage, equipment reliability, and the environment. Liquid cooling is an increasingly desirable, or even a necessary, alternative where massive computing power is required, such as in AI factories.

---

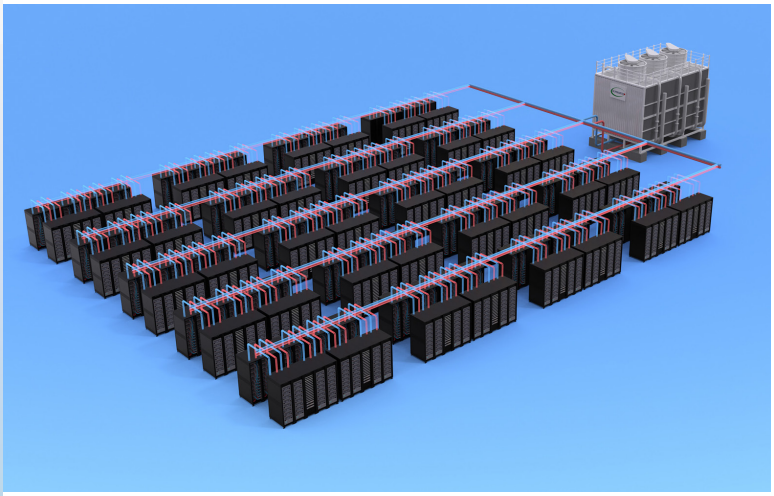
1 Key Trends and Technologies Impacting Data Centers in 2024 and Beyond, AFCOM, March 14, 2024

2 Rack Density is Rising, The Uptime Institute, December 7, 2020

White paper

The benefits of liquid cooling are becoming clearer, particularly in new data center construction. AFCOM's survey found that 40% of data center managers plan to use the technology to accommodate rack density increases. Saying liquid cooling is "set to go mainstream," Dell'Oro Group forecasts that the market could reach $15 billion in annual revenue over the next five years.[3]

The most common type of liquid cooling — direct-to-chip cooling — pipes cool/cold liquid through hoses to plates on top of heat-generating components such as CPUs, graphics processing units (GPUs), and memory cards. The liquid absorbs the heat from the components, and then the resulting warm liquid circulates to a cooling device or a heat exchanger. The cooled fluid is then circulated back to the cold plates.



Other liquid cooling types include immersion cooling, rear-door heat exchangers, and in-row cooling. Each has its advantages and trade-offs, with the choice depending on factors such as server density, power usage, cooling capacity, and data center design. Direct-to-chip cooling is likely to be the most popular technology for the foreseeable future due to its ability to cool the liquid and be selectively applied to new racks.

6 reasons you need liquid cooling

Liquid cooling has many advantages over traditional air cooling. Here are some of the most important:

1. Lower cooling infrastructure costs

The cost of building a data center is somewhat linked to the amount of

3 Data Center Liquid Cooling Market Set to Go Mainstream and Top $15 B over the Next Five Years, According to Dell'Oro Group, Dell'Oro Group, July 23, 2024

power that needs to be delivered to the entire data center, mainly for the IT equipment inside. Air-cooled systems draw in vast amounts of external air, filter it, cool it, and circulate it throughout the facility. Each stage requires power. Air conditioning equipment takes up space, requires maintenance, and needs backup power sources such as uninterruptible power supplies.

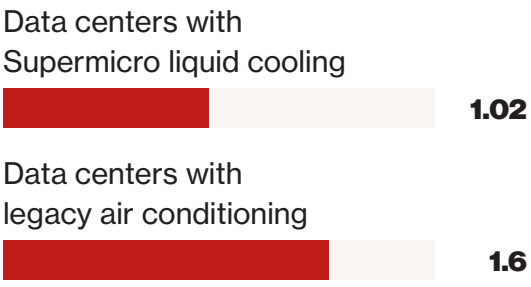
Liquid cooling dramatically reduces the need for bulky air conditioning systems and the power needed to run them. It also minimizes energy waste, by applying cooling directly to GPUs and CPUs. That, in turn, lowers the overall power requirements of the facility. There are also ancillary benefits. Data centers that consume less power are less taxing on local utilities, encounter fewer zoning issues, and enable more computing power or draw less power while fitting within a power budget that the local utility may impose.

2. Lower PUE

Power usage effectiveness (PUE) is a metric derived when the total energy consumed by a data center is divided by the energy used to run the IT equipment. A rating of 1.0 is considered optimal efficiency.

Supermicro's liquid cooling implementations have achieved near-perfect PUEs of as low as 1.02, compared to 1.6 for legacy air-cooled data centers. Reduced power requirements mean that liquid cooling can lower overall electricity costs by up to 40%. Not only are smaller power infrastructure and cooling systems required but the reduction in cooling-related megawatts also translates to savings on construction materials and space, as fewer physical components must be installed.

Figure 1 | **PUE Rating**



The operational expense (OPEX) in data centers is significantly

influenced by electricity usage, which is directly related to the need for cooling. McKinsey estimates that cooling accounts for about 40% of the average data center's energy consumption.[4]

Liquid cooling requires far less electricity than air conditioning because liquids absorb and transfer heat up to 1,000 times more efficiently than air. Thus, nearly all the energy the data center consumes goes to running the IT equipment. The result is that cooling-related energy consumption is up to 90% less than with air-cooling.[5]

3. Increased compute density

In addition to its impressive cost-saving benefits, liquid cooling enables data center operators to pack computing elements more tightly, thus enhancing space usage. The efficiencies can't be achieved with conventional air cooling, which requires physically larger servers that can move more air to permit adequate airflow. That limits how many machines can be housed in a single rack.

Liquid cooling removes this restriction. Because cooling is applied directly to components, more servers can be placed in a single rack, fewer racks are needed, and there is no need for large heat sinks that take up significant volume in the server.

Some data centers can double computer density by using smaller, liquid-cooled servers that provide the same — or even more — computing power as large ones. Moving from an 8U air-cooled AI server to a 4U liquid-cooled server doubles computational power density. This has major cost implications since data centers cost $600 and $1,100 per square foot.[6] Having fewer racks and less infrastructure also translates into less equipment to repair and maintain.

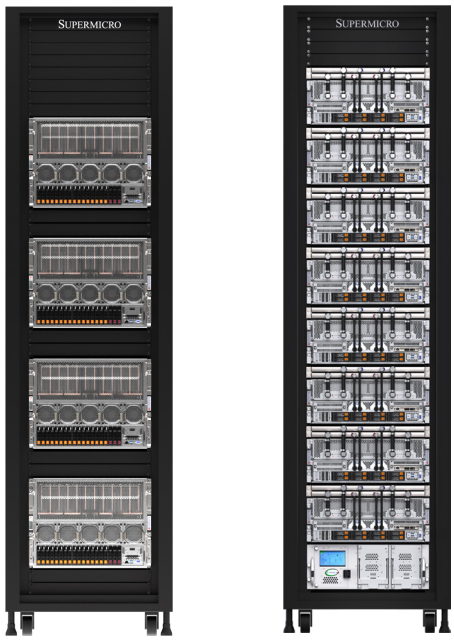
---

4 Investing in the rising data center economy, McKinsey & Company, January 17, 2023

5 Reducing Carbon Footprint with Liquid Immersion Cooling, Intel and Green Revolution Computing, April 2022

6 How Much Does it Cost to Build a Data Center? Dgtl Infra, November 5, 2023

Figure 2 | **A rack of 8U air-cooled servers and a rack of 4U liquid-cooled servers**



4. Better performance

One of the most exciting performance-related benefits of liquid cooling is its ability to maintain higher processor performance levels for extended periods.

Many processors, especially in HPC and AI workloads, have a "turbo boosting" feature. This technology enables the processor to temporarily increase its clock speed above its base operating rate when the workload demands it. However, this boost can be short-lived if temperatures climb and heat threatens the integrity of the electronics.

Liquid cooling allows for more sustained turbo boost periods, enabling processors to run at their boosted clock speeds longer or indefinitely without thermal throttling. Some tests have shown that this longer boost period can improve performance on HPC applications by as much as 13%.

5. Workplace benefits

A less tangible but important advantage of liquid cooling is a reduction in noise levels of up to 55% in data center environments. The powerful fans needed to cool air in servers in conventional data centers can generate noise levels as high as 96 decibels (dBA), which is louder than a subway train.[7]

7 Why Data Centers Are Loud, and How to Quiet Them Down, Data Center Knowledge, June 2, 2023

Studies have shown that noise levels of 95 dBA can put hearing at risk in less than an hour.[8] They have also found that high noise levels increase worker stress, reduce productivity, increase worker compensation claims, and cause a greater instance of worker absences and sick leave.

Liquid cooling systems operate silently or with minimal noise. They contribute to a quieter workplace that enhances communication and concentration for staff working in the facility.

6. Lower carbon footprint

For all of its exciting potential, AI is expected to have a troubling environmental impact. The demand for AI and HPC processing is expected to drive a compound annual data center growth rate of 40.5% through 2027. Data center energy consumption is forecast to grow even faster — with a CAGR of 44.7% — more than doubling between 2023 and 2028.[9]

Since fossil fuels generate 82% of U.S. energy,[10] AI will inevitably increase carbon emissions.

Liquid cooling offers a path toward more sustainable operations. It requires far less power, resulting in lower electricity consumption and a reduced carbon footprint with fossil fuels as the source. Liquid cooling also offers the opportunity to reuse waste heat for district heating or other thermal energy reuse applications, such as greenhouses.

For organizations aiming to meet sustainability goals, liquid cooling provides an opportunity to cut energy use and dependence on fossil fuels. This supports corporate social

8 Workplace Noise: More than just "All Ears", National Institute for Occupational Safety and Health Science Blog, June 28, 2018

9 IDC Report Reveals AI-Driven Growth in Datacenter Energy Consumption, Predicts Surge in Datacenter Facility Spending Amid Rising Electricity Costs, International Data Corp., September 24, 2024

10 U.S. Renewable Energy Factsheet, Center for Sustainable Systems, University of Michigan, 2024

White paper

responsibility initiatives, helping companies meet green energy targets and reduce their environmental impact.

Preparing for the future

Liquid cooling will become increasingly necessary for new and top-of-the-line hardware. Accelerated Computing giant NVIDIA recommends using liquid cooling for its next generation of GPUs, NVIDIA Blackwell, rated at up to 1,200 watts.[11] The reference specification for the NVIDIA GB200 NVL72 system electromechanical design, which the company recently donated to the Open Computing Project, encompasses direct-to-chip liquid cooling technology to accommodate its 36 NVIDIA Grace CPUs and 72 NVIDIA Blackwell GPUs and is 25X more energy efficient than its predecessor.

By implementing liquid cooling today, data centers can stay ahead of technological advancements and remain equipped for future demands. With the increasing power of next-gen high-performance GPU chips, air cooling is no longer efficient or adequate. Liquid cooling provides a more robust, efficient, and cost-effective alternative to traditional air cooling solutions.

From reducing construction and operational costs to enabling higher computing density, improving environmental sustainability, and boosting performance, liquid cooling's many advantages are apparent. The shift toward liquid cooling isn't just a trend — it's a necessary evolution for advancing AI and high-performance computing.

Learn about Supermicro's liquid cooling solution.

11 NVIDIA's Fully-Enabled Blackwell B200 GPUs Consume Up To 1200W, Completely Different Architecture From Hopper, WCCFTech, March 22, 2024