

STULZ

CLIMATE. CUSTOMIZED.

Liquid Cooling

**When performance meets sustainability –
the data center cooling challenge**



Contents

1.	Data centers in the age of artificial intelligence (AI)	3
2.	Liquid cooling and the drivers for adoption	3
3.	Liquid cooling technologies explained	4
	Direct-to-chip liquid cooling	4
	Immersion liquid cooling	5
4.	Design and implementation considerations	6
5.	Energy efficiency and sustainability benefits	8
	Calculation examples: CDU plate heat exchanger optimization	9
	Case 1: Standard high-density data center CDU	10
	Case 2: The trade-off between efficiency and size	11
	Case 3: Efficiency without expansion – optimizing TCS return temperature	12
6.	Economic analysis	13
7.	Future trends and developments	14
8.	Summary	14

1. **Data centers in the age of artificial intelligence (AI)**

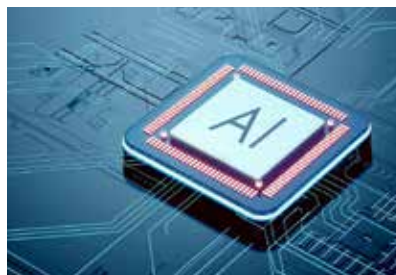
Data center operators are currently facing a significant challenge. The rise of AI is forcing the need for greater computing power and this, in turn, is leading to higher energy consumption. At the same time, new sustainability-based regulations and rising customer expectations are emphasizing the need for more sustainable operations. This puts data center operators in a predicament – how can they address the technological demands being placed on them without compromising on performance or sustainability?

2. **Liquid cooling and the drivers for adoption**

Nvidia, with its range of high-performance graphics processing unit (GPU) architecture such as the Nvidia Blackwell, is pushing data center power densities from 70 kW per IT rack to 120 kW. Likewise, with ever more powerful servers, GPUs and central processing units (CPUs) often consuming over 300 W – 400 W of power, future projections indicate these levels could double or even triple within the next 5 – 7 years. Intel envisions a future where chip density increases tenfold, potentially enabling the integration of a trillion transistors within a single chip package by 2030.

This type of immense computing power results in high temperatures being concentrated in small areas, something that conventional air cooling can no longer deal with effectively. Put simply, air, as a cooling medium, has reached its physical limits since it cannot store or transfer enough heat. In contrast, liquid is up to 1,000 times more efficient as a cooling medium. Therefore, liquid cooling, once considered a 'nice to have' solution, is rapidly becoming essential in today's high-performance computing landscape.

The rapid evolution of AI, and the technology needed to facilitate it, inevitably intersects with the data center industry's sustainability goals and net-zero targets. As a result, as well as a need for innovative and energy efficient cooling solutions, data center operators are facing new challenges from an ecological perspective, with greater sustainability now required.



To illustrate this point, the Energy Efficiency Directive was adopted by the European Commission in September 2023 and includes a requirement to monitor the energy performance of data centers. This involves the creation of a European Union (EU) wide database to collect and publish relevant data. Data centers operating in the EU with an installed information technology power demand of at least 500 kW must report key performance indicators to a European database for sustainability assessment.

There are various country specific regulations in place too, such as the German Energy Efficiency Act, which came into force in October 2023. This sets specific targets for energy consumption effectiveness and stipulates obligatory PUE targets for data centers based on when they first began operation. A data center operating before 1st July 2026 will need to have a PUE metric of 1.5 from July 2027 and 1.3 from July 2030. For facilities operating from 1st July 2026, the PUE target will be 1.2.

This all results in a dilemma – performance and sustainability are required in equal measure but are contradictory in terms of implementation.

3. **Liquid cooling technologies explained**

Liquid cooling helps tackle the thermal challenges associated with AI, improving the overall efficiency and sustainability of data centers. With inadequate cooling temperatures, servers, CPUs and GPUs cannot deliver optimal performance. However, liquid cooling can handle exceptionally high densities by bringing cooling mechanisms closer to technology, providing a heat transfer capacity that far surpasses traditional air cooling. As a result, liquid cooling is now firmly on the radars of forward thinking data center owners and operators.

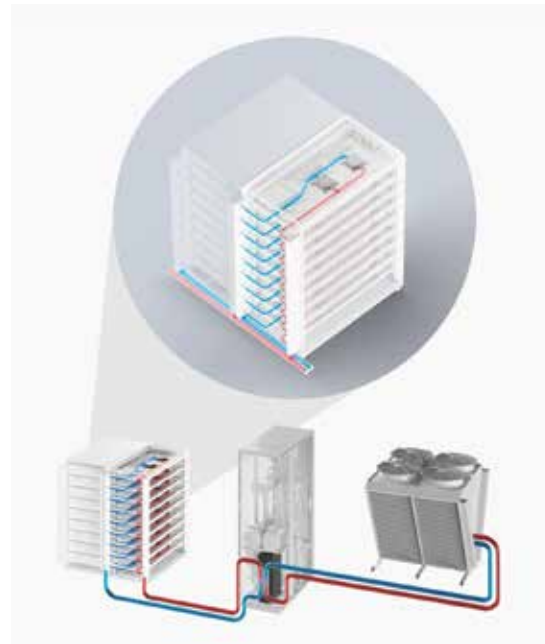
There are two primary types of liquid cooling available, which differ in design and efficiency. Let's look at each of them in more detail:

Direct-to-chip liquid cooling

Direct-to-chip liquid cooling systems use a variety of coolants such as water, water/glycol mixtures, deionized water or dielectric fluids on an oil basis. With high thermal conductivity and low electrical conductivity, a fluid is circulated through a cold plate attached to a server, CPU or GPU. As the liquid flows, heat is transferred from the hot processor to the cooler liquid and removed. Due to its high specific heat capacity, the cooling liquid is far more effective than air at dissipating heat, helping maintain optimal processor temperatures, regardless of workload or external conditions.

Direct-to-chip liquid cooling removes 70–90 % of the heat generated by rack equipment through water, with the remaining 10–30 % dissipated through air cooling technology. This method is highly effective in heat transfer, allowing it to support higher server, CPU and GPU densities, while also providing significant heat reuse potential.

Direct-to-chip cooling can be integrated into existing designs, minimizing disruption to operations and streamlining the deployment process. Liquid cooling systems generally consume less energy than air cooling systems, resulting in reduced operational costs and a smaller carbon footprint for data centers. With direct-to-chip cooling, higher heat loads can be implemented per rack, which means that less space is required to process the same data center heat load. It also means that as the data center design becomes different, savings become possible due to the size of the building, and also in the operation of the cooling system.



Immersion liquid cooling

Immersion liquid cooling involves servers, GPUs and CPUs being submerged into a thermally conductive dielectric liquid or coolant. Heat is extracted from the system by circulating the liquid, which comes into direct contact with hot components and then passes through cooling heat exchangers. Fluids used for immersion cooling are specifically chosen for their excellent insulating properties, ensuring they can safely interact with energized electronic components.

In single-phase immersion cooling, heat is directly transferred from the immersed components to the surrounding liquid. The fluid is then cycled out of the immersion tank by a coolant pump that runs through a heat exchanger and is returned to the immersion tank at a lower temperature, where it continues this heat transfer cycle.

In two-phase immersion cooling, heat from immersed server components causes the fluid to become gaseous. There is a condenser in the upper part of the tank, which is cooled from outside by a water circuit. The fluid cools down upon reaching the condenser, transitions into a liquid state, and flows back down to absorb heat once more.

Immersion liquid cooling can remove almost all the heat generated by the electronic components and can significantly reduce or eliminate water usage from the cooling system design. However, while immersion cooling is highly efficient there are considerations such as its high initial capital expenditure (CAPEX) and ongoing maintenance, as well as the disposal of dielectric liquid. Also, it still requires around five per cent air cooling due to heat loss in the tank and not all IT equipment can integrate immersion cooling.

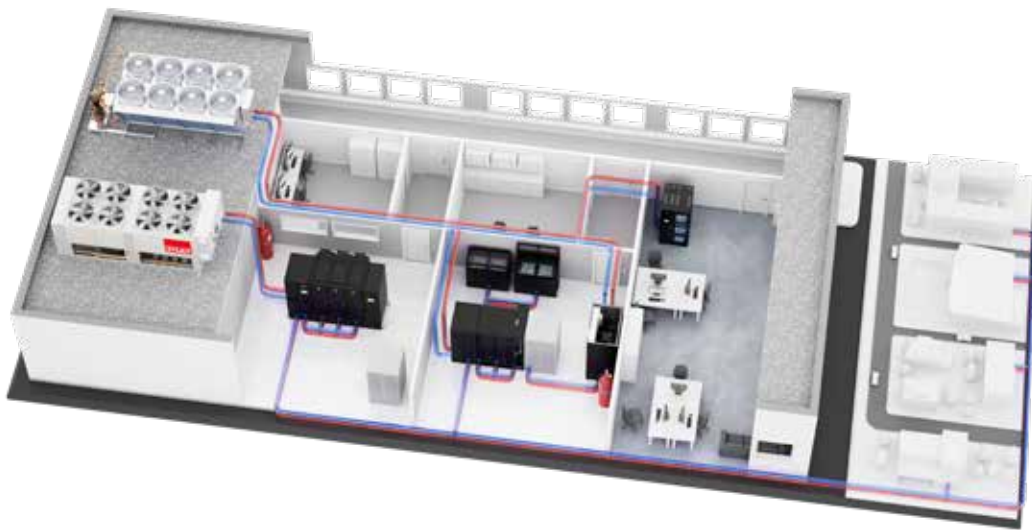


4. Design and implementation considerations

IT equipment is extremely sensitive and must be kept within a strict temperature range to perform to its full potential. Straying from these limits means technology will be in an environment it was not designed for, which will shorten its lifespan. Any reduction in lifespan will increase costs and mission-critical equipment that ceases to operate correctly can wreak havoc on a data center.

In a rapidly evolving digital landscape, scalability is essential and liquid cooling not only addresses current demands but also ensures data centers are future proof. By increasing density, liquid cooling optimizes the use of available space, making it possible to support resource intensive AI. However, liquid cooling requires a holistic approach that integrates advanced monitoring systems, physical redundancy and automation.

Deploying liquid cooling is a major undertaking that demands thorough planning and attention to various factors including the facility's existing layout, current thermal management approach, workloads and budget. Those new to deploying liquid cooling must address key considerations such as establishing an efficient cold water supply and hot water removal, understanding the role of a cooling distribution unit (CDU) and determining whether it is necessary, and identifying the appropriate heat rejection architecture for the system. Conducting a total cost of ownership (TCO) analysis is highly recommended to optimize operations by replacing outdated or inefficient equipment, thereby reducing operational expenditure (OPEX).



Existing infrastructure should be evaluated to determine its adaptability for more power intensive AI workloads. Physical spaces should be assessed to confirm if any existing raised floors can support the combined weight of new power and cooling systems, as well as to identify viable access routes for piping installations. Additionally, a thorough inspection of the facility should be conducted to address any maintenance needs for the existing infrastructure, as issues like contamination or degradation in pipes or equipment could cause inefficiencies or failures. Furthermore, the on-site water supply must be reviewed to ensure it meets the requirements for the planned liquid cooling system.

Although there is currently a lack of defined industry standards when it comes to designing a liquid cooled data center, several key factors must be addressed to ensure proper installation of components like CDUs, cold plates, tanks and piping. To convert an air-cooled system to liquid cooled system, in most cases a new rack and servers will be required. This is due to the space required for the internal piping of the cold plates and the manifold, which may also require additional holes for pipe exits.

Installing a cutting-edge CDU allows for precise control over each side of the liquid cooling system, facilitating better management of coolant flow rates, temperatures and pressure, which boosts system efficiency. Leading CDUs also feature sanitary grade stainless steel pipelines, alongside a range of structural, electrical and control options including the flexibility to accommodate different liquids and power loads .

When converting to liquid cooling, a building's structure must be carefully evaluated. Compared to air conditioning systems, the use of DCLC increases the weight per rack. However, with immersion cooling, particularly in larger facilities, there is a risk of substantially increased point loads on the ceiling due to higher server densities and the volume of fluid involved. To address this, structural engineers must assess the ceiling's load bearing capacity before the conversion.



The CDU operates based on fixed FWS (Facility Water System) values and TCS (Technology Cooling System) outlet temperatures, while the return temperature from the servers varies depending on computing power fluctuations.

5. **Energy efficiency and sustainability benefits**

The data center sector is responsible for a remarkable share of total global electricity consumption, making it an important focus of political, economic and ecological consideration. This high energy demand has prompted growing pressure on data center operators to optimize their facilities and contribute to the push for net-zero emissions. As a result, many are committing to carbon neutral targets, emphasizing transparency and accountability in their environmental impact.

In a typical data center, a significant proportion of its total energy consumption is used by the cooling system. The rapid increase in energy consumption due to the growing demands of AI workloads, combined with rising power costs and stricter government regulations, highlights the critical need to reduce data center energy usage.

Sustainability has become a key factor in vendor selection, evolving from a perceived cost to a fundamental component of business strategy. However, a conflict is already emerging in today's high density data center infrastructure that will become even more pronounced in the coming years. It will require a solution – this is the challenge faced in designing a data center that is as powerful as possible and at the same time as energy efficient as possible.

Liquid cooling offers a more effective solution for data centers but a technical conflict lies in the juxtaposition of the lowest possible approach temperature (APT) and the lowest possible temperatures on the server side, with the smallest possible temperature spread across the servers themselves. The lowest possible APT should increase energy efficiency and reduce OPEX by minimizing the amount of heat energy that the chiller must remove. And the chiller is precisely the component with the highest power consumption and the highest OPEX in the overall cooling circuit.

+ APT (Approach Temperature)

There are four temperature values in a plate heat exchanger - two on each side of the FWS and TCS. APT refers to the temperature difference between the FWS inlet and the TCS outlet.

The ability to precisely adjust liquid cooling systems to meet the specific cooling requirements of each rack helps also to minimize energy waste. Simultaneously, any excess heat generated can be efficiently reused and redirected, as liquid cooling systems can support temperatures up to 60 °C, while heat pumps can elevate these to the 75 °C needed for integration into district heating systems.

However, another clearly recognizable trend is that water temperatures and media are being primarily defined by the performance requirements of the chip manufacturers. Only then is an attempt made to optimize the CDU and the chiller for greater sustainability. Practical examples show temperature requirements of 35/25 °C or 36/26 °C in the TCS when using water in the FWS and PG25 (propylene glycol 25 %) in the TCS – especially in greenfield projects. In brownfield applications, on the other hand, we see values of 20/30 °C in the FWS with water and PG25 in the TCS, combined with the lowest possible APT.

Calculation examples: CDU plate heat exchanger optimization

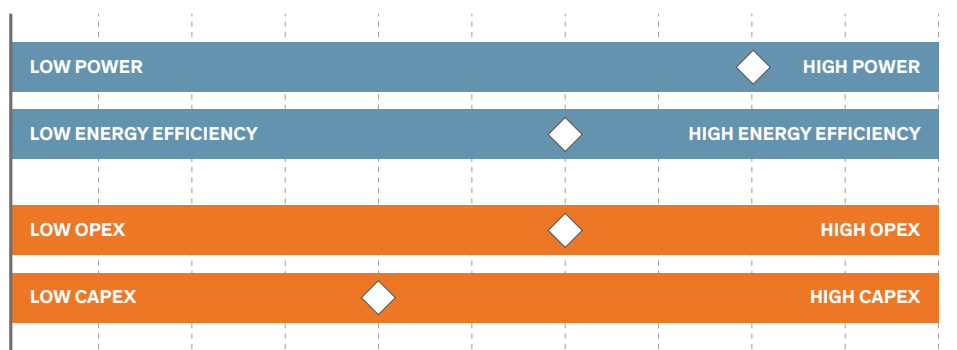
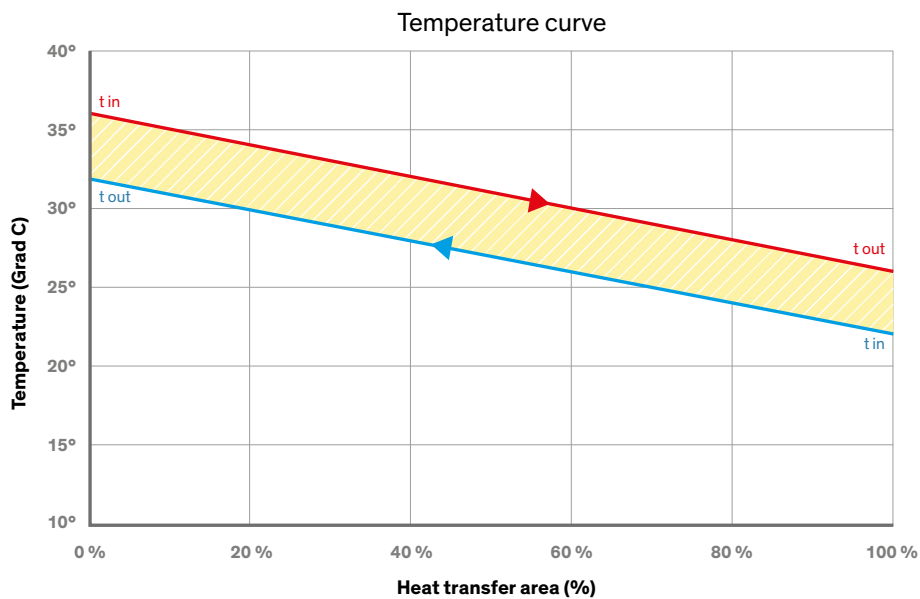
When designing a data center cooling system that balances both performance and energy efficiency, data center planners and operators must consider various thermodynamic factors. The following calculations illustrate three possible approaches to selecting a Cooling Distribution Unit (CDU) plate heat exchanger, each with different design parameters and implications for efficiency and cost.

Case 1: Standard high-density data center CDU

In a conventional high-density data center setup, a CDU with a cooling capacity of 1,000 kW is designed with the following operating parameters. The Facility Water System (FWS) operates at 22 °C/32 °C, while the Technology Cooling System (TCS) runs at 36 °C/26 °C. This design results in an Approach Temperature (APT) of 4 K and requires a total heat transfer area of 71.8 m². This configuration represents a standard selection where the heat exchanger efficiently transfers heat between the FWS and TCS while maintaining a reasonable APT.

However, as energy efficiency becomes a more critical factor, alternative approaches must be evaluated.

Case 1			
Cooling capacity	kW	1,000	
APT	K	4	
		FWS	TCS
Volume flow (V)	m ³ /h	23.9	25.3
Inlet temperature	T _{in}	22	36
Outlet temperature	T _{out}	32	26
Temperature difference (ΔT)	K	10	10
Mean log. Temperature difference		4	
Medium		Water	PG25
Specific heat capacity	c	4.183	3.95
Heat exchange area (A)	m ²	71.8	



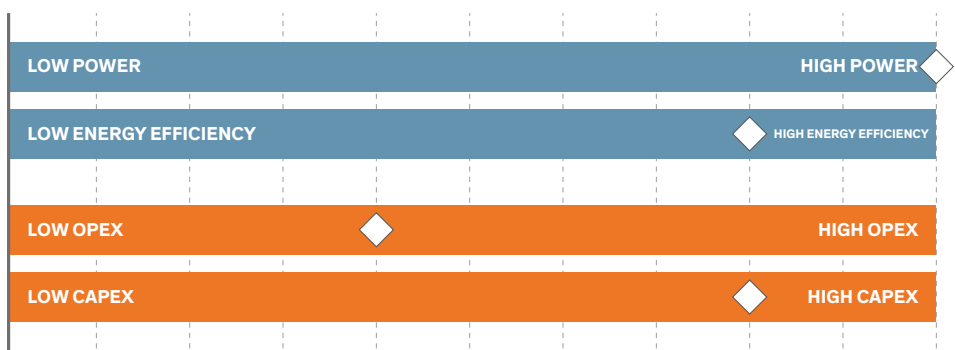
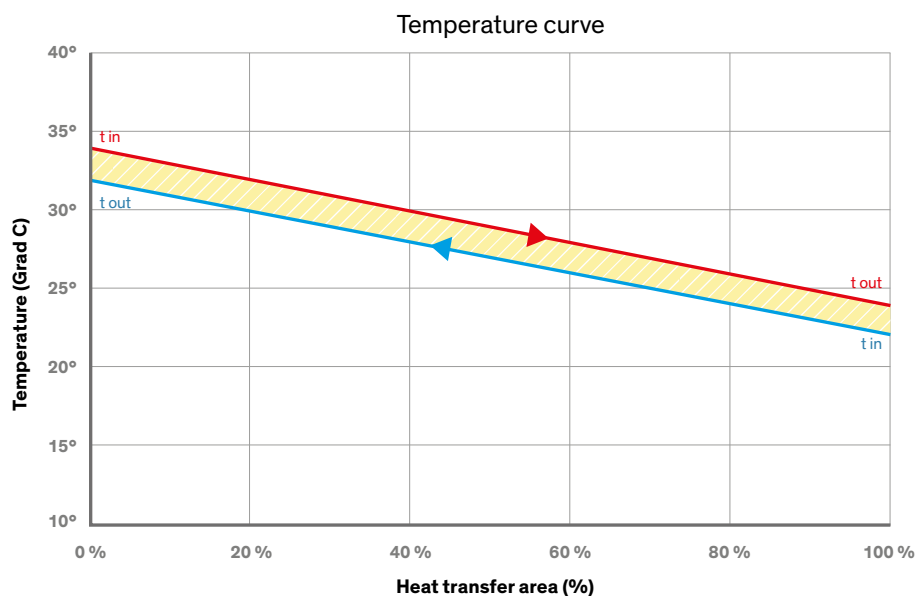
Case 2: The trade-off between efficiency and size

In this scenario, the FWS temperatures remain the same, and the temperature difference on both sides is maintained at 10 K. The APT is reduced to 2 K to enhance efficiency.

As a result, the total required heat transfer area must be doubled, leading to a heat transfer area of 143.6 m². This option benefits AI-driven data centers where workloads fluctuate significantly, with servers shifting between low-load operation and overclocking.

The increased heat transfer area ensures that high-volume flow on the TCS side and lower temperatures are available at all times. However, this approach requires higher capital expenditure (CAPEX) for the larger heat exchanger and modifications to the data center infrastructure.

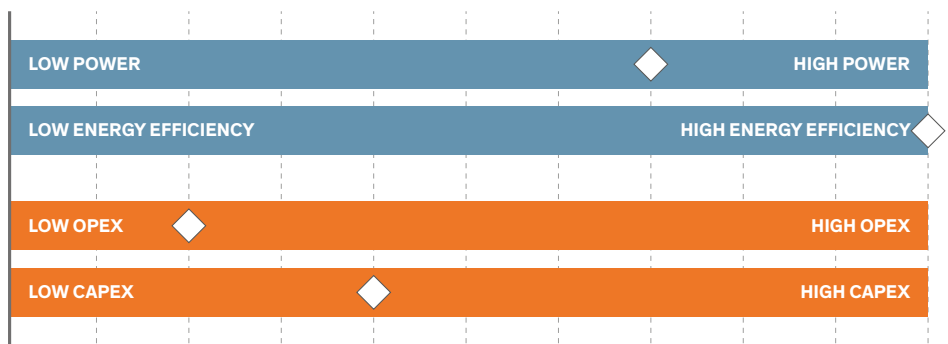
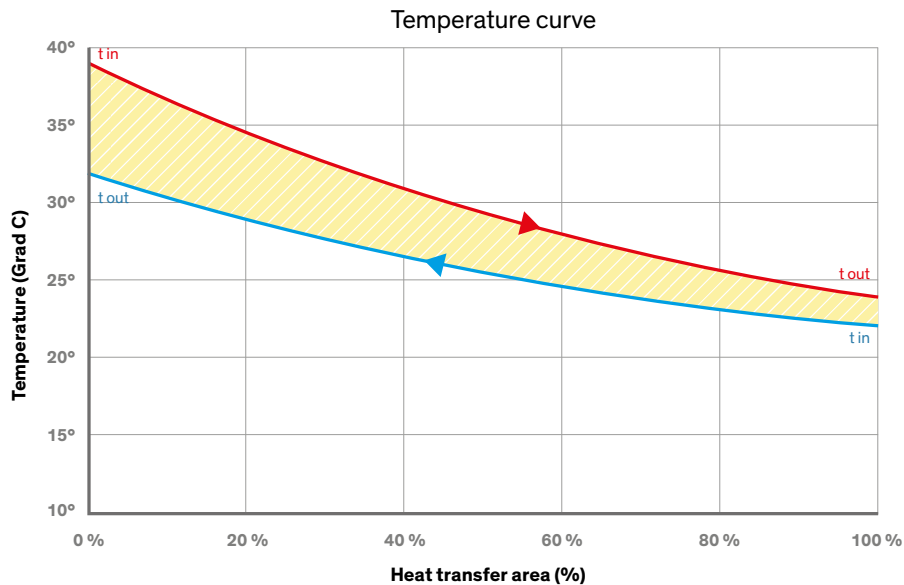
Case 2			
Cooling capacity	kW	1,000	
APT	K	2	
		FWS	TCS
Volume flow (V)	m ³ /h	23.9	25.3
Inlet temperature	T _{in}	22	34
Outlet temperature	T _{out}	32	24
Temperature difference (ΔT)	K	10	10
Mean log. Temperature difference		2	
Medium		Water	PG25
Specific heat capacity	c	4.183	3.95
Heat exchange area (A)	m ²	143.6	



Case 3: Efficiency without expansion – optimizing TCS return temperature

An alternative method to maximize energy efficiency without doubling the heat exchanger size is to increase the TCS return water temperature while keeping the APT at 2 K. This adjustment ensures that the mean logarithmic temperature difference ($\Delta T_{m,log}$) remains at 4 K, allowing the system to achieve the required cooling capacity without increasing the heat exchanger area and the CAPEX. This results in a heat transfer area of 71.8 m², the same as in Case 1. This approach prioritizes lower operational expenditure (OPEX) due to a reduced TCS flow rate by 33 %, which in turn lowers pump power consumption. The downside is the potential need to evaluate IT equipment capabilities to handle a wider temperature spread across servers and ensure the system can respond quickly to computing power fluctuations.

Case 3			
Cooling capacity	kW	1,000	
APT	K	2	
		FWS	TCS
Volume flow (V)	m ³ /h	23.9	16.9
Inlet temperature	T _{in}	22	39
Outlet temperature	T _{out}	32	24
Temperature difference (ΔT)	K	10	15
Mean log. Temperature difference		4	
Medium		Water	PG25
Specific heat capacity	c	4.183	3.95
Heat exchange area (A)	m ²	71.8	



6. Economic analysis

Liquid cooling CAPEX includes the procurement and installation of specialized equipment such as cold plates, immersion tanks or heat exchangers. Additionally, retrofitting legacy data centers to accommodate liquid cooling can involve structural changes and adjustments to existing infrastructure, further driving up CAPEX. According to industry estimates, liquid cooling can cost up to 30–50 % more than traditional air cooling during initial deployment.

While the initial costs may be high, liquid cooling offers compelling long-term savings and return on investment. It can reduce energy consumption for cooling by 30–40 % compared to air cooled systems. It can also help to achieve PUE values as low as 1.1, significantly improving operational efficiency. Ultimately, supporting compute intensive applications on an industrial scale must focus on building high-performance infrastructure systems that are precisely engineered and fit for purpose now and in the future.





7. **Future trends and developments**

According to MarketsandMarkets, the global data center liquid cooling market is expected to increase from \$ 4.9 bn in 2024 to \$ 21.3 bn by 2030, representing a compound annual growth rate (CAGR) of 27.6 %. Furthermore, the results of a 2024 survey from The Register found that more than a third of respondents (38.3 %) expect to employ some form of liquid cooling infrastructure in their data centers by 2026, up from just 20.1 % having done so when the research was initially conducted.

It's clear that AI is revolutionizing digital infrastructure, creating unprecedented demand for computational power and introducing significant challenges for data center operators. As the computational demands of AI continue to grow, CPUs, GPUs and servers will evolve and become even more powerful, so understanding the benefits of liquid cooling technologies is now essential.



8. **Summary**

To navigate this complex landscape, operators must carefully evaluate and implement strategies that align with technological requirements, regulatory demands and the shift toward sustainable infrastructure. Data center operators will, however, continue to fall into two distinct camps – those who focus on maximum performance and those who prioritize sustainability issues. It is currently clear that it is not possible to achieve an optimum in performance and an optimum in sustainability at the same time. The operator will have to decide on the right mix. Technically, each variant can be designed and can be offered by the liquid cooling technology manufacturers – at different costs.

The future will show how these two important factors can be harmoniously combined to create the next generation of data centers that are able to adapt to the rapid developments in chip technology.



ABOUT THE AUTHOR

Jonathan Aaron Jürgens is a Junior Product Manager at STULZ GmbH, where he has been shaping the future of data center cooling since 2020. Holding a bachelor's degree in industrial engineering, Jonathan plays a key role in the development and market implementation of the CyberCool CMU. With hands-on experience from over 80 global liquid cooling projects since the CMU's launch, Jonathan has established himself as the in-house expert on liquid cooling infrastructures. To further share this expertise, he conducts liquid cooling seminars worldwide, equipping data center stakeholders with the knowledge needed to optimize cooling performance while enhancing sustainability.

ABOUT STULZ GMBH

STULZ is a global provider of innovative, mission-critical cooling solutions. Founded in 1947 and with its headquarters in Hamburg, Germany, it is represented by 24 subsidiaries, a network of more than 150 partners and nine production sites worldwide. Its solutions provide precise temperature and humidity control, as well as reliable cooling capacity, to safeguard the operation and longevity for data centers, telecom infrastructure and industry. STULZ prioritises energy efficiency and sustainability in its product development and manufacturing processes to deliver future proof cooling solutions.

For further information, please visit our website at www.stulz.com

CLOSE TO YOU AROUND THE WORLD

STULZ Company Headquarters

STULZ GmbH

Holsteiner Chaussee 283
22457 Hamburg
Tel. +49 40 5585-0
products@stulz.de

For further information, please visit our
website at **www.stulz.com** or download our
STULZ Products and Services app.

AUSTRALIA
AUSTRIA
BELGIUM
BRAZIL
CHINA
FRANCE
GERMANY
INDIA
INDONESIA
IRELAND
ITALY
MALAYSIA
MEXICO
NETHERLANDS
NEW ZEALAND
POLAND
SINGAPORE
SOUTH AFRICA
SPAIN
SWEDEN
THAILAND
UNITED KINGDOM
USA

**+ 150 PARTNERS
WORLDWIDE**