



Disrupting Data Centre Design

How liquid cooling is
reshaping data centre
designs and cost models



kpmg.ie/datacentres2030

Liquid cooling will be one of the key enablers in AI computing architectures being deployed by data centre developers. But what is it, and under what circumstances does it impact the design of data centres, their costs, and return on investment?

This report examines the transformative potential of liquid cooling, an emerging technology that is poised to become a cornerstone of modern data centre design. We will explore the diverse approaches to liquid cooling, its operational advantages, and how it impacts data centre configurations. Additionally, this report provides insights into the cost modelling implications, offering a thorough perspective of how liquid cooling can optimize both performance and expenditures for data centre developers.

What are the various approaches to cooling?

Air cooling

Air cooling, which supports up to approximately 70kW per rack, has long been the de facto standard for data centres. However, this approach is now falling out of favour. New data centres are increasingly moving away from air cooling as their primary method, opting for alternative liquid-to air or liquid-to-liquid based solutions. Similarly, many brownfield sites with aging and outdated infrastructure are also shifting towards liquid cooling to enhance efficiency and meet modern demands.

Liquid cooling

Liquid cooling comes in various forms, but it's important to understand that liquid cooling is not a single product. It is a system and an ecosystem comprising various components such as Coolant Distribution Units (CDUs), cold plates, manifolds, liquid-cooled servers, heat rejection units, and complementary air-cooling components. Most vendors are unveiling product roadmaps that include hybrid (liquid-air cooling), liquid-liquid cooling, and immersion cooling.

1. Liquid-to-Air (L2A) cooling

Liquid-to-air cooling is a significant area of development and may be suitable for existing data centres that lack the infrastructure, space, or investment for liquid-to-liquid systems. This hybrid approach, which is currently the most common form of cooling utilised in data centres supporting cloud and edge computing, serves as a stop-gap or bridge solution. It allows data centres to improve cooling efficiency without the need for extensive retrofitting. By combining elements of both air and liquid cooling, it offers an incremental step towards fully liquid-based systems while managing costs and minimising disruption.

2. Liquid-to-liquid (L2L) cooling

2.a) Direct-to-chip liquid cooling involves attaching cold plates directly to heat-generating components like CPUs and GPUs. The liquid coolant flows through these plates, absorbing heat from the components. This method supports densities of up to approximately 120kW/rack (such as the Nvidia GB200 NVL72, the company's most powerful Blackwell rack-scale design) and can allow for somewhat higher densities before reaching its limit. Direct-to-chip liquid cooling is currently experiencing the fastest adoption and has emerged as the predominant technology because it is easier to transition to from a layout perspective.

2.b) Immersion cooling can support even higher densities, above 120kW/rack, although the transition is more challenging. In immersion cooling, servers or components are fully submerged in a fluid that absorbs heat directly from the components. There is a consensus among hyperscalers about the potential of immersion cooling, a concept that smaller players like crypto miners have already adopted. However, immersion cooling has not yet been scaled to large data centres.

While many in the industry see immersion cooling as the ultimate solution, the most efficient currently commercialised technology is direct-to-chip liquid cooling.



Why liquid cooling now?

Air cooling has become expensive due to the high volumes of air needed, as air is not a very good conductor.

The trend towards high-density computing and AI requires significantly more power per rack, resulting in more heat. Increased heat raises the risk of equipment in data centres malfunctioning or catching fire. The challenges of using air cooling are exacerbated by the introduction of newer chips with higher heat loads.

Consequently, the need for liquid cooling becomes more apparent. Liquid cooling allows for more precise targeting of cooling areas and is generally more efficient. With newer chips that can only be efficiently cooled by liquid methods, liquid cooling is experiencing a resurgence. If data centre operators and their customers want to use the leading-edge chips, they must consider liquid cooling

At HPE's AI Day on October 10 2024, it was indicated that the value of liquid-cooled servers will reach USD 35 billion by 2027, up from USD 5 billion in 2023, with a 50% CAGR from 2023 to 2027. Hyperscalers are the main adopters, contributing over 50%, followed by tier-2 and tier-3 service providers.¹

Will liquid cooling be the de-facto choice for new data centre designs?

For now, air and liquid cooling will coexist. Liquid cooling systems will not be standalone in the foreseeable future; they will be supplemented by traditional air-cooling technologies. This combination results in various ratios, such as 60:40 or 70:30 (Liquid:Air), depending on the specific solution. Direct-to-chip solutions can use liquid cooling for up to 80-90% of heat removal, but air cooling remains necessary for the approximately 20% of heat not captured by the cold plate.

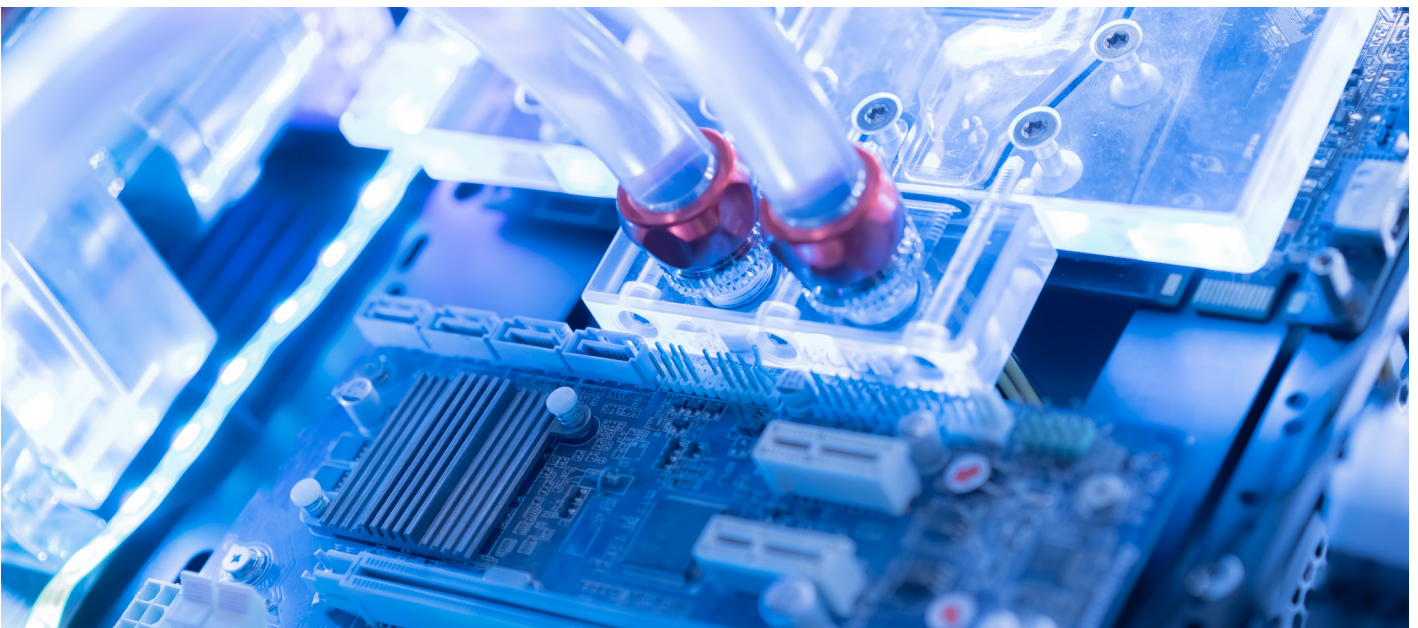
In retail colocation facilities, where customers make the decisions, a variety of cooling solutions will also be seen under one roof.

Are there standard solutions?

Various liquid cooling solutions are emerging as vendors adopt a fast-paced, innovative approach. There are new players contesting the space. However, established cooling suppliers (e.g. Vertiv, CoolIT, Motivair) stand out due to their scale and service, which is more critical than price when managing risks associated with these innovative systems being rolled out at scale for the first time.

Efforts are being made to make liquid cooling systems modular and scalable. Achieving standardisation in terms of temperatures, reasonable flow rates, and standard connections is crucial; currently, these parameters are unique to specific vendors. Actions taken by large OEMs/ODMs are expected to drive this standardisation.

¹ Key takeaways from HPE's AI Day", October 11, 2024 [\[Link\]](#)



What is the impact on new data centre designs?

1: New future-proof reference designs

Reference designs for data centres remained stable for a significant period. By adopting uniform designs across multiple projects, large data centre developers (such as hyperscalers and larger colocation providers) reduced procurement complexity, facilitated easier maintenance, and improved scalability.

However, the design of data centres has recently become volatile. Existing reference designs are no longer adequate and need to be replaced. The industry is at an inflection point with AI on the cusp of significant growth. To future-proof data centres, operators now need to plan for racks in the 100-200kW range.

Designers face the challenge of developing flexible plans that allow air and liquid cooling solutions to coexist, potentially transitioning to more sophisticated liquid cooling approaches as cooling technology evolves and immersion cooling becomes more commercially viable over time.

This includes incorporating provisional space for new items of owner furnished equipment such as Coolant Distribution Units (CDUs), thermal storage buffer tanks (to allow for the accumulation of sufficient chilled water during a chiller restart as a result of a power loss) and additional mechanical Uninterruptible Power Supply systems (UPSs) and batteries to support the CDUs or mechanical UPSs to support the Chillers. Commissioning of liquid-to-liquid data modules adds a further level of complexity to an already onerous process so commissioning plans must include these detailed requirements from day one and be integrated into purchase requirements and tender documents at the earliest opportunity to mitigate supply chain challenges.

2: Design Considerations

Some general key considerations for transitioning from air to liquid cooling that data centre designers will need to consider are:

2.a) Implementation Options

- **UPS-backed chillers:**

This method involves adding a UPS for each chiller to negate any chilled water temperature losses in the time between a utility power loss and the time it takes for the chillers to restart under generator load as the UPS will allow the chillers to continue to run under normal conditions.

- **Thermal Energy Storage Tanks (TES):**

This method allows for the inclusion of TES tanks or buffer vessels on the chilled water system. Chilled water is stored in the tanks which provides a “backup” supply of chilled water to allow the circuit to continue to perform optimally in the time it takes for the generator to restore load to the chillers.

2.b) Space Optimisation

Introducing liquid-to-liquid cooling optimises space utilisation within the data module by allowing for a higher IT capacity and greater rack densification. The space optimisation and the rack density within the data module are entirely dependent on the design stage of the project and can be classified into two categories:

- **Non-Densified:**

This scenario is reserved for projects currently on site and under construction that must be retrofitted, and projects whose design is too far advanced to a point where it is not possible to alter building permits without causing significant programme delays. In a non-densified scenario, the physical space within the data module will remain the same, as will the IT load. The space within the data module that could be utilised for racks is taken up by CDU galleries and in some instances UPS batteries and TES tanks.

- **Densified:**

This scenario is used early in the design process and is more suited to projects at feasibility stage as it gives design teams more flexibility to ensure adequate consideration of for the inclusion of any new plant associated with L2L. In a densified scenario the physical space within each individual data module will remain the same, however the IT load of each individual data module will double. This allows developers to double the IT capacity of the building (power dependant) or potentially reduce the building footprint by 35-40% while maintaining the same IT load. However, this reduction will depend on various considerations during the permitting and planning phase, such as whether certain plant equipment can be placed on the roof.

2.c) Mechanical and electrical considerations (non-exhaustive)

Category	Consideration	Details
Mechanical Considerations	CRAHs (computer room air handlers) / Fan Walls	Approx. 30% reduction in the quantity of CRAHs (liquid to air) with the implementation of CDUs (coolant distribution units) (liquid-to-liquid)
	Chilled water (CHW) thermal energy storage tanks (TES) tanks	As per design requirements, depending on the chosen option.
	Pipework connections	Additional connections to and from building risers and CHW headers to connect CDUs on the primary side.
	Data module chilled water pipework	Also known as the Technical Cooling System (TCS) on the secondary side of the CDUs, should be stainless steel.
	Pressure Independent Linear Flow Valves (PILFVs)	Additional linear pressure independent linear flow valves will be required above each server rack which are controlled and monitored by the Building Management System (BMS)
	Tech loop	It will be essential to evaluate how the tech loop is load tested under various conditions and handed over to data center ensure proper function, reliability.
	Water leak detection & drip trays	Additional drip trays and water leak detection will be required for pipework over the racks.
Electrical Considerations	Low voltage panels and cables	New Low Voltage Sub-Distribution Panels to supply mechanical UPSs and CDUs
	Primary & secondary busbar systems	Primary and secondary busbar configurations will change to accommodate densification of server racks in the data modules.
	Control wiring and structured cabling	Additional cabling will be required to allow valves to be controlled and monitored by the BMS.

It should also be noted that in AI/HPC environments, liquid cooling systems must always be powered on to manage the high-performance components. In air-cooled systems, airflow can be adjusted as needed, depending on internal and external factors. CPU-based data centres had higher tolerances for temperature fluctuations. According to ASHRAE, traditional data centres operated within 18°C to 27°C, allowing for minor delays in cooling without major impacts.

However, AI workloads, which rely on GPUs, cause significant spikes in heat output that require immediate cooling. The Uptime Institute reports that GPUs generate rapid temperature spikes, and NVIDIA guidelines highlight that inadequate cooling can lead to thermal throttling and hardware failures. Therefore, new data centre designs must ensure continuous operation of liquid cooling systems, integrating UPS backed power for pumps and CDUs – posing new mechanical and electrical considerations for data centre designers.

3: Total cost of ownership (TCO) of the design

Liquid cooling technology is a relatively new approach in data centre cooling compared to traditional air-cooled designs, and it typically uses fewer commercial off-the-shelf components. Designers and integrators need to help customers understand the TCO and the efficiency gains associated with liquid cooling systems.

While there may be higher upfront costs due to specialised equipment and developing an appropriate electrical topology, the total capital expenditure can be comparable to or even lower than air-cooled systems when considering the entire setup. Additionally, liquid cooling systems tend to result in lower overall power usage and can operate at higher temperatures, enhancing their efficiency.

L2L Solution	Average € per MW uplift *	Consideration
<p>Option A: Reduce building footprint by 35-40%</p>	<p>€400k to €600k per MW</p>	<p>Data centres early in the design phase If the development project is still in the feasibility stage, there is still time to optimize the design of the new buildings. For example, the space requirements for a densified L2L data module are the same as that of a liquid-to-air cloud data module, but the IT load for densified L2L is double that of the cloud. So, depending on power availability, the IT capacity can be doubled.</p> <p>The cost estimate for reducing the building footprint by 35-40% is €400 - €600k per MW.</p> <p>Additionally, the cost estimate includes the TES tech loop and commissioning, which would otherwise fall into the customer fit-out scope at €900k - €1.1m per MW. By absorbing these costs, the developer is providing a credit to the customer, as these expenses would typically be the customer’s responsibility.</p>
<p>Option B: Densification</p>	<p>€1.2m – €1.5m per MW</p>	<p>Data centres late in the design phase If the data center design is advanced (meaning various permits are in place and certain specific grid power load constraints have been assigned) and the building structure cannot be altered, the racks can still be densified, however, only half of the white space or the IT space will be utilized within the building. This fallow space in the “spare” data halls can then be used to house the TES tanks and UPS batteries. Overall, this use of space is somewhat sub-optimal.</p> <p>Again, the cost estimate includes the TES tech loop and commissioning, which would otherwise fall into the customer fit-out scope at €900k - €1.1m per MW.</p>

* Cost considerations are based on a TES tank option with primary and secondary busbar as the preferred feeder for the IT load and are construction costs only. Furthermore, these costs are based only on General Contractor/Owner-Furnished Contractor-Installed expenses and do not include any adjustments for the actual racks, which are direct customer items.

At KPMG Ireland, we understand the critical role data centres play in today's digital economy. Our Data Centre Ecosystem Hub is dedicated to providing comprehensive advisory services that span the entire data centre lifecycle.



Christopher Brown

Partner
Head of Strategy
KPMG in Ireland
t: +353 87 050 4453
e: christopher.brown@kpmg.ie



Eoin Dunphy

Director
Construction Advisory
KPMG in Ireland
t: +353 87 606 1628
e: eoin.dunphy@kpmg.ie



Morgan Mulooly

Associate Director
Strategy
KPMG in Ireland
t: +353 87 050 4330
e: morgan.mulooly@kpmg.ie



Veronica Guedes

Senior MEP Cost Manager
Construction Advisory
KPMG in Ireland
t: +353 87 166 7807
e: veronica.guedes@kpmg.ie



David Quinn

Senior MEP Cost Manager
Construction Advisory
KPMG in Ireland
t: +353871667807
e: david.john.quinn@kpmg.ie



Stephen O'Grady

Senior MEP Cost Manager
Construction Advisory
KPMG in Ireland
t: +353871116990
e: stephen.ogrady@kpmg.ie



© 2024 KPMG, an Irish partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved.

The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavour to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

The KPMG name and logo are registered trademarks of KPMG International Limited ("KPMG International"), a private English company limited by guarantee.

If you've received this communication directly from KPMG, it is because we hold your name and company details for the purpose of keeping you informed on a range of business issues and the services we provide. If you would like us to delete this information from our records and would prefer not to receive any further updates from us please contact unsubscribe@kpmg.ie.

Produced by: KPMG's Creative Services. Publication Date: December 2024 (11014)