INTELLIGENCE UPDATE

Liquid cooling will not outgrow its high-density niche

Jacqueline Davis 21 Jan 2026

Deployment of direct liquid cooling (DLC, as cold plate or immersion systems) remains overwhelmingly concentrated in applications where air cooling is no longer a practical alternative. As a wave of AI infrastructure was being built, DLC manufacturers invested their engineering resources in maximizing cooling performance for high-performance IT — namely AI training hardware. By comparison, DLC manufacturers have done little to address the unresolved operational issues that hinder DLC adoption for IT applications with differing priorities.

Despite its benefits in performance and efficiency, DLC is unlikely to achieve true mainstream adoption in the next 3 years. Specifically, most operators will not deploy DLC at scale for their mission-critical applications. These applications lack an imperative to densify, and remaining difficulties in resiliency design and data center operations severely undermine the business case for DLC.

Even as DLC manufacturers scale up (or are acquired by large data center vendors), their resources to tailor subsequent generations of cooling products remain finite. For as long as the buildout of high-density AI training data centers continues, it will steer DLC design priorities away from the needs of high-resiliency (typically lower-density) business applications.

DLC expands, but within bounds

In 2026, DLC remains confined to the “niche” domain of supercomputing and similar applications — although the rise of generative-AI training has significantly expanded this domain and established DLC’s presence in many more facilities.

Generative AI training is essentially a form of supercomputing, and it is their common traits that make generative AI a natural fit for DLC. AI training jobs run to completion without real-time user input, similarly to batch computing. Failure of individual compute nodes or even restarting training usually do not incur an immediate financial loss, because revenue typically comes later from inference applications (see *AI embraces liquid cooling, but enterprise IT is slow to follow*).

Compared to AI training and high-performance computing (HPC) hardware, IT that supports

© COPYRIGHT 2026 UPTIME INSTITUTE. ALL RIGHTS RESERVED.

business-critical applications (e.g., databases) typically has lower thermal density and a higher resiliency objective. Here, DLC adoption faces a twofold challenge: high standards of resiliency are typically more complex and costly to achieve with liquid cooling, and mainstream IT applications often lack an urgent need to densify.

Maintaining application uptime through events such as human error, planned maintenance, and equipment failure typically relies on redundant cooling equipment. Air cooling equipment uses the facility air as both a thermal reservoir and a shared interface for redundant air conditioners or air handlers. In contrast, DLC closely couples the facility infrastructure with the IT (e.g., with immersion baths or tubing connections to cold plates). Plumbing and valves to connect redundant CDUs or manifolds need to be carefully engineered and operated, and switching between units can cause troublesome turbulence effects that are difficult to model and predict.

The volume of the liquid coolant is also much less than that of the air in the data hall, affording less time for standby generators to start and synchronize before taking over the cooling load following a power outage. Acceptable ride-through time typically requires UPS power for CDU pumps, at a minimum — and it is still uncommon for facility equipment to be backed by UPS power. The industry also lacks a singular consensus on the division of ownership and maintenance for DLC components between facilities teams and IT teams (see *Hold the line: liquid cooling's division of labor*).

The above issues are a persistent concern for data center operators, even in applications where densified IT is necessary for performance, such as in HPC. Early in the planning process when placing an application and its IT, organizations weigh the benefits and drawbacks of densified IT. Less dense IT is more likely to be accommodated in existing data hall space and usually results in more predictable and familiar infrastructure design conditions in new data center developments.

Many mission-critical applications that serve user requests in real time do not demonstrate a clear return on investment from deploying high-powered processors and densely packed racks. Running these applications using IT with more conventional (lower) processor power and work capacity minimizes the portion of the workload that is affected by a single server failure due to cooling or other issues. Such IT often tends toward lower utilization to leave margin for peak demand. Serving these applications with high utilization and high resiliency on powerful liquid-cooled IT would incur equipment and engineering costs for redundant DLC, or a software resiliency alternative.

Investment reinforces existing priorities

Rising demand for conventional, lower-density data center space has persisted alongside the generative AI boom, yet it has received far less attention by comparison. In the next 3 years and beyond, data center operators will build both high-density space for AI training and more conventional data halls to host business-critical applications. Overall, investment in generative AI is unlikely to detract from finance for conventional data centers, because they have proven business models and solid expectations for further growth.

© COPYRIGHT 2026 UPTIME INSTITUTE. ALL RIGHTS RESERVED.

However, the engineering and manufacturing capacities of DLC manufacturers are finite and, for the next few years, they are likely to remain focused on the fast-growing AI and HPC applications that align with DLC's present design priorities. Generative AI applications generate sufficient demand for liquid cooling to keep order books full for the major DLC manufacturers. As they continue to scale their manufacturing capacities, DLC manufacturers have more to gain by continually improving their products' maximum cooling performance and suitability for this sector because it represents the bulk of their revenue today.

By comparison, re-engineering DLC to solve the complex problems of daily operation to a high standard of resiliency appears to be a difficult proposition, with a considerably less compelling business case. Even if DLC manufacturers devote engineering resources to facilitate the design and operation of redundant liquid cooling systems — likely at higher cost to end users— it is unclear whether operators of mission-critical IT would embrace it. This is especially true where there is little appetite to densify IT.

Manufacturers of DLC will thrive and post record sales in the next few years. However, DLC adoption across a broader cross-section of the industry will remain elusive until mission-critical applications develop a stronger motive than thermal density.

In the long term, DLC's energy efficiency potential could become more important — especially where grid constraints incentivize maximizing the proportion of power used for IT, or where legislation steps in to shape this decision-making process. IT densification is not predestined, and DLC manufacturers may want to take a long view and work toward a more favorable cost/benefit profile for future applications. The buildout for generative AI training will not escalate indefinitely, and a trained model's speed to market will not retain the primacy it has today. Some organizations (especially those performing incremental training on top of a pre-trained foundation model) will likely perform training on lower-density clusters, accepting some limitations to model size or training speed. The early-2025 announcement of DeepSeek, trained on lower-density hardware yet delivering competitive results, suggested that further software innovation could provide an alternative to runaway densification.

The Uptime Intelligence View

DLC now has more deployment and investment than ever — but still serves a targeted subset of IT. Supercomputing and HPC drove the development of today's DLC designs, and AI training provided a ready opportunity to grow deployment with little need to change design priorities. Overall, data center investment and buildout appear almost limitless, but DLC manufacturing capacity and engineering resources are finite. Rack density and high heat flux remain the strongest drivers for DLC by far — and until a broad variety of applications develop stronger drivers for DLC, true mainstream adoption will remain elusive.

Other related reports published by Uptime Institute include:

© COPYRIGHT 2026 UPTIME INSTITUTE. ALL RIGHTS RESERVED.

AI embraces liquid cooling, but enterprise IT is slow to follow
Hold the line: liquid cooling's division of labor

ABOUT THE AUTHOR

Jacqueline Davis

21 Jan 2026

Jacqueline is a Research Analyst at Uptime Institute covering global trends and technologies that underpin critical digital infrastructure. Her background includes environmental monitoring and data interpretation in the environmental compliance and health and safety fields.

jdavis@uptimeinstitute.com

© COPYRIGHT 2026 UPTIME INSTITUTE. ALL RIGHTS RESERVED.

About Uptime Institute

Uptime Institute is the Global Digital Infrastructure Authority. Its Tier Standard is the IT industry's most trusted and adopted global standard for the proper design, construction, and operation of data centers – the backbone of the digital economy. For over 25 years, the company has served as the standard for data center reliability, sustainability, and efficiency, providing customers assurance that their digital infrastructure can perform at a level that is consistent with their business needs across a wide array of operating conditions.

With its data center Tier Standard & Certifications, Management & Operations reviews, broad range of related risk and performance assessments, and accredited educational curriculum completed by over 10,000 data center professionals, Uptime Institute has helped thousands of companies, in over 100 countries to optimize critical IT assets while managing costs, resources, and efficiency.

© COPYRIGHT 2026 UPTIME INSTITUTE. ALL RIGHTS RESERVED.