PATHFINDER REPORT

451 Research | Advisory

The Tide is Turning Toward Liquid Cooling in Datacenters

How immersion cooling helps next-generation applications while reducing risks

JULY 2018

COMMISSIONED BY

GRC GREEN REVOLUTION COOLING
The Immersion Cooling Authority

About this paper

A Pathfinder paper navigates decision-makers through the issues surrounding a specific technology or business case, explores the business value of adoption, and recommends the range of considerations and concrete next steps in the decision-making process.

About 451 Research

451 Research is a leading information technology research and advisory company focusing on technology innovation and market disruption. More than 100 analysts and consultants provide essential insight to more than 1,000 client organizations globally through a combination of syndicated research and data, advisory and go-to-market services, and live events.

NEW YORK
1411 Broadway
New York NY 10018
+1 212 505 3030

SAN FRANCISCO
140 Geary Street
San Francisco, CA 94108
+1 415 989 1555

LONDON
Paxton House
(Ground floor)
30, Artillery Lane
London, E1 7LS, UK
P +44 (0) 207 426 1050

BOSTON
75-101 Federal Street
5th Floor
Boston, MA 02110
Phone: +1 617.598.7200
Fax: +1 617.357.7495

Executive Summary

Direct liquid cooling (DLC), a set of techniques that removes heat from chips and other electronics without air as a medium, has evolved from niche applications in mainframes, high-performance systems and gaming rigs to become a datacenter-class technology. Unlike past efforts that were specific to a brand of servers or a large supercomputing project, the DLC systems available today are flexible and can support a wide range of systems, including custom specifications.

Contrary to previous expectations, it will not be high-density racks that trigger most upgrades to DLC systems at mainstream operators, but underlying trends in processor and storage technology. Over the last 10 years, processors – not only general-purpose microprocessors but also graphical processing units (GPUs) and other accelerators – have become increasingly power hungry as chip designers push toward the frontiers of silicon performance. Chip frequencies have also become dependent on cooling capacity. In short, more cooling capacity means more performance for the dollar, something that will become more obvious and difficult for a growing number of infrastructure operators to ignore. Also, storage has become compatible with DLC, and a fully liquid-cooled data hall is now a practical possibility.

Even though 451 Research believes that much of DLC adoption – outside of high-performance computing – will be tactical at first in support of specific applications, a longer-term plan for phased facility-wide installation is possible now, and a growing number of operators (and tenants in commercial facilities) will see it as desirable. The rollout of DLC as a component of their technology roadmap will support efforts to cut operational costs while increasing infrastructure availability and reliability via simplified maintenance and lower failure rates.

Key Findings

- Direct liquid cooling, a well-understood and mature set of techniques, is making a comeback in datacenters to solve the various challenges and limitations of air systems.
- DLC is quickly becoming a relevant consideration for enterprises and service providers, moving beyond the classic supercomputing use case.
- Datacenter managers will find that most facilities will struggle handling next-generation IT systems for performance-critical applications.
- DLC is not only for compute anymore; storage has become a prime candidate for liquid cooling to improve density, availability and reliability.

Haven't we been here before? The déjá vu of liquid cooling

What exactly is direct liquid cooling, and why is it markedly different? The key point is that heat sources, in this case electronics such as processors and memory, transfer their thermal energy via contact with the coolant or a liquid-cooled heat sink. There is no air involved in this heat-exchange process: the coolant transfers the captured thermal energy away from IT systems for heat rejection. This is important to differentiate from forms of cooling where liquid cools the air entering or leaving the IT system, even if closely coupled to the IT – typical examples of that are front- and rear-door heat exchangers.

The argument for DLC is simple: liquids are orders of magnitude better at heat transfer than air, and because there are no fans, DLC is less noisy and creates much less vibration. For these very reasons, the technology has been in use for many decades. The first DLC systems appeared in the 1960s with IBM's System/360 mainframes. In fact, at one point, the technology provided the primary form of cooling for datacenter IT when liquid-cooled mainframes dominated the server market up to the 1990s, mostly running on bipolar semiconductor technology. Evidently, vendors and customers alike trusted liquid cooling, a tested and understood technique, with the most mission-critical applications.

What happened to DLC then? The advent of mass-market computers based on complementary metal-oxide semiconductor (CMOS) technology happened. The relatively inexpensive and low-power server systems – built around CMOS chips – that gained the upper hand in datacenters by the second half of the 1990s used heat sinks and fans on processors to drive down the cost per system with little regard to facility build and operational costs. Because the server population was a tiny fraction of what it is today and chips were much less power hungry, there was no need for such cost concerns. Equally important was the fact that the speed of semiconductors, not power, limited performance. At the time, those in the server market decided DLC was an unjustified cost.

This switch to more power-efficient CMOS technology established air cooling as the norm in datacenters. Liquid-cooled electronics in the datacenter were relegated to exotic use cases such as mainframes and some high-performance computing clusters, which remains the prevalent view among datacenter operators. Massive investments in air-cooling systems by vendors and operators have created major inertia in the industry that resists change.

However, 451 Research maintains a positive view of DLC. We believe that change is not only happening but is inevitable – the question for operators is when and how, not if. For the most part, the move toward DLC will be gradual – tactical at first, as opposed to a strategic wholesale move away from air. Driving this shift are multiple underlying technology trends that we discuss in detail below to explain why we think DLC will go mainstream this time, but the shift is also a product of development efforts and a growing installed base that resulted in well-honed, datacenter-class DLC systems with a considerable operational track record.

LIQUID COOLING EXPLAINED

There are several approaches to DLC in datacenters, but the main ones are:

- **Cold plates –** A system of heat sinks on major electronic components with coolant circulating inside them. Several DLC systems use water as the coolant; others use refrigerants. The benefit of water is that it is a good conductor of heat, and it is inexpensive and plentiful. Refrigerants are non-corrosive and may require lower volumes and less pumping at low pressures, so there is no risk from leakage. Cold plates are also relatively easy to retrofit to conventional servers, but this technology doesn't capture 100% of the load – there remains a need for air systems.
- **Immersive –** Server systems are submerged in a form of liquid, such as engineered dielectric fluids (also called synthetic), mineral oils and their blends. These systems have several advantages over cold plates, chief of which is that all components can be directly immersed in these liquids. This improves thermal efficiencies and often results in simplified DLC systems with no requirement for additional air-based cooling for the immersion units.

Vendors have developed a fair number of cold-plate and immersion systems that differ in engineering choices to find the sweet spot in the market. Other approaches to DLC that haven't been widely adopted include the use of liquid Freon, liquid nitrogen and liquid helium as coolants – more typical of high-performance servers and workstations that operate chips beyond factory specifications. Use of these substances in datacenter DLC remains rare, and their cost and volatile nature mean that they are unlikely to be used at scale in datacenters. IBM and others have also developed processors equipped with microfluidic channels used to carry coolants, and possibly even power, directly inside chips, but these remain largely experimental.

Removing heat from IT components and capturing it in a liquid is only half of the DLC process; a variety of technologies are also used to eject that captured heat from the datacenter. For example, suppliers have shown the potential of warm-water cooling, where the inlet water to the DLC system does not require any mechanical cooling. In these examples, methods such as dry coolers and heat exchangers (liquid-to-liquid or liquid-to-air) are used to lower the temperature of the liquid before it reenters the IT system. When coupled with chillers at an added cost, cooling performance can be increased via the delivery of low-temperature fluids to the chips. However, these considerations for cost and performance, albeit important, are secondary to the benefits DLC brings to datacenter operations.

Why This Time is Different

Although liquid cooling has been around for decades, wider adoption has failed to materialize despite repeated calls from industry participants, so why would it happen now? Simplified facilities engineering and 10-20% reduction in capital costs haven't convinced mainstream operators to uproot existing design and operational practices. Crucially, the adoption of DLC as standard in data halls would have reduced configuration flexibility to support all kinds of IT systems – facilities teams couldn't have possibly forced the hands of IT infrastructure managers. Also, DLC has traditionally been applicable only to compute, while storage systems (hard disk drive arrays and magnetic tapes) needed air cooling.

On the other hand, IT infrastructure teams outside the domain of high-performance computing clusters have rarely asked facilities to support DLC. They didn't have to: power densities, contrary to expectations, remained moderate in the 4-6kW range of average cooling requirement per cabinet. Recent datacenter builds, optimized using computational fluid dynamics simulations, have typically no problem handling the odd 10-20kW cabinets, more typical of recent converged and hyperconverged systems at large enterprises and IT service providers.

Yet 451 Research believes wider adoption of DLC will start happening in the next 12-24 months. Even though facilities costs and rack densities, the classical arguments in favor of DLC, have not converted most of the datacenter industry to liquid cooling, there are some major factors at play that will direct much more attention to it.

BETTER COOLING MEANS FASTER APPLICATIONS

PROCESSOR DESIGN EVOLUTION

First is the evolution of processor design. In the past, processor performance was largely static in the sense that a nominal frequency was fixed in the factory, and vendors such as Intel, AMD and NVIDIA guaranteed that their chips could reach and sustain those speeds if thermal conditions were met. This picture has become much more dynamic in recent years. Nominal frequencies as guaranteed minimums remain; however, most processors are now opportunistic in their performance within an electrical and thermal envelope. When a modern chip is below its critical temperature and power limits but one or more of its cores are highly loaded, it will start to increase frequencies (and voltage levels in support) in steps to meet the need for performance.

This will continue until the chip hits a thermal or electrical limit. For a few cores, it will likely be factory caps for maximum frequency the vendor sets for reliability reasons. But for many cores, which is typical for highly parallel datacenter workloads, it will be to prevent thermal runaway. If the chip begins to close in on critical temperature, it will slow back down, even throttle if needed. In this modern clocking regime, unlike in the static world of the past, more cooling capacity translates to more performance, sometimes considerably more: most high-performance many-core chips tend to be thermally, and not electronically, limited (propagation delay in their circuitry).

What DLC systems can do is support higher levels of sustainable processor performance and help clear spikes in the application load, maintaining acceptable response times. Chip vendors today allow customers to configure (within limits) their processors' power and thermal limits to match them better with their cooling and operational objectives. Large customers that buy chips by the thousands or tens of thousands will likely be able to ask for processors that are factory-configured for even higher power levels – in some cases 50% higher – when DLC is used. Anecdotally, some operators achieved 15-25% speedups as a result when compared to the fastest available stock (air-cooled) processor. This is significant in light of the very modest gains in core performance in recent years, and should improve not just overall capacity of a given infrastructure, but responsiveness too.

This is only the start to the trend, 451 Research predicts. Chip vendors still design their processors assuming the use of air cooling. This will likely change in a few years, not only because some major customers will start asking for chips, in large quantities, that can utilize the higher cooling capacity and stable temperatures provided by DLC systems, but also because of the general pressure to deliver ever higher levels of performance from a single piece of silicon. Appetite for performance comes not only from large computational problems such as deep learning and analytics against big datasets that are becoming common at major enterprises and service providers, but also from more generic infrastructures. Examples include the need to rein in on infrastructure sprawl (densification and overall economics), as well as to boost performance-critical enterprise applications such as in-memory computing and database engines that support revenue generation.

SEMICONDUCTOR TECHNOLOGY TRENDS

This leads us to our next point about silicon: semiconductor technology trends. It is the nature of CMOS technology scaling that the speed at which transistors shrink outpaces reductions in their power use. In response, chip vendors have invested in designing more electronically optimized circuitry and highly parallel chips with more moderately clocked cores. Over the long term, however, there is no escaping the underlying trend that a square inch of silicon gets higher power with every generation.

If the IT industry wants to take full advantage of even larger scales of integration (and if history is any guide, it will in order to drive better IT infrastructure economics), allowances for much higher power consumption per chip must be made. Notably, Intel's latest-generation Xeon processors have abandoned the previous limit of 165 watts (thermal power) per chip and introduced a 205-watt class for some of its server chips with 24-28 cores. It is a fair assumption that some large buyers will prefer that power class for parts of their infrastructure.

Even though air systems (heat sinks and fans) can still cool 205 watts, it is clear that they are reaching their practical boundaries and will become unviable for much higher-powered parts. Some AMD and NVIDIA GPUs have already passed the 200-watt threshold for some time now and are approaching 300 watts. It is no surprise that most GPU-heavy systems, even gaming PCs, use heat pipes to spread the heat to a larger surface area. Signaling the direction of travel, Intel has publicly said that a higher-power version of its 28-core chips will come later in 2018 in limited quantities that will sustain much (up to 50%) higher clock speeds when low-temperature liquid cooling is used. While this is more of a technology demo than a viable product, it demonstrated the level of performance that will be possible with next-generation chips and DLC systems.

STORAGE HAS BECOME A PRIME CANDIDATE FOR LIQUID COOLING

After silicon, the second key component to greater DLC adoption is storage technology. Hard disk drives (HDDs) and tape libraries meant that in the past, even if an operator wanted to adopt DLC, air systems had to stay in place, which canceled out some of the cost benefits. Wholesale transition to DLC was not even possible, which meant DLC was added complexity for unclear benefits bar HPC clusters.

This is not the case anymore. Tapes are largely gone from most data halls, having largely lost their position as online backup systems for primary storage and now largely fulfill disaster recovery and archiving roles. HDDs, while still widely used in servers despite fast-paced adoption of flash for primary storage, have undergone a key change in their build: all high-capacity models today are filled with helium. Manufacturers Seagate, Toshiba and Western Digital use helium, which has a much lower density than air, to reduce friction and turbulence when platters rotate at high speeds. This reduces power requirements, heat generation and, most importantly, lets the read-write head find and keep position with higher precision. A beneficial side effect for DLC systems is that helium-filled drives are hermetically sealed, which means they are compatible with full immersion cooling techniques.

With solid-state drives (SSDs) and helium-filled HDDs, a fully liquid-cooled datacenter is not a distant aspiration anymore but a practical possibility. This changes the long-term outlook from one where air cooling and DLC have to live side-by-side indefinitely to one where a datacenter operator can craft a technology roadmap that leads to predominant or exclusive use of DLC systems in the whitespace. The key consideration here is lowered complexity to the baseline of air systems, which means less expensive design and engineering fees, as well as simplified maintenance procedures: fewer components mean less chance of failures that lead to service disruption.

Perhaps more important is the effect on IT operations: DLC should also mean fewer IT component failures, chiefly drives. HDDs are susceptible to high temperatures (a burden that forces data halls to remain cool), and even more so to humidity (mandating humidity controls). Although a seemingly risky move, submerging HDDs into liquid will likely result in fewer drive failures. Developments in storage technology have not only enabled the use of DLC, but will favor it once customers realize and quantify these benefits.

Higher infrastructure performance and better reliability at ultimately lower cost when compared to an air system will attract an increasing number of customers to DLC systems, 451 Research believes.

Use Cases

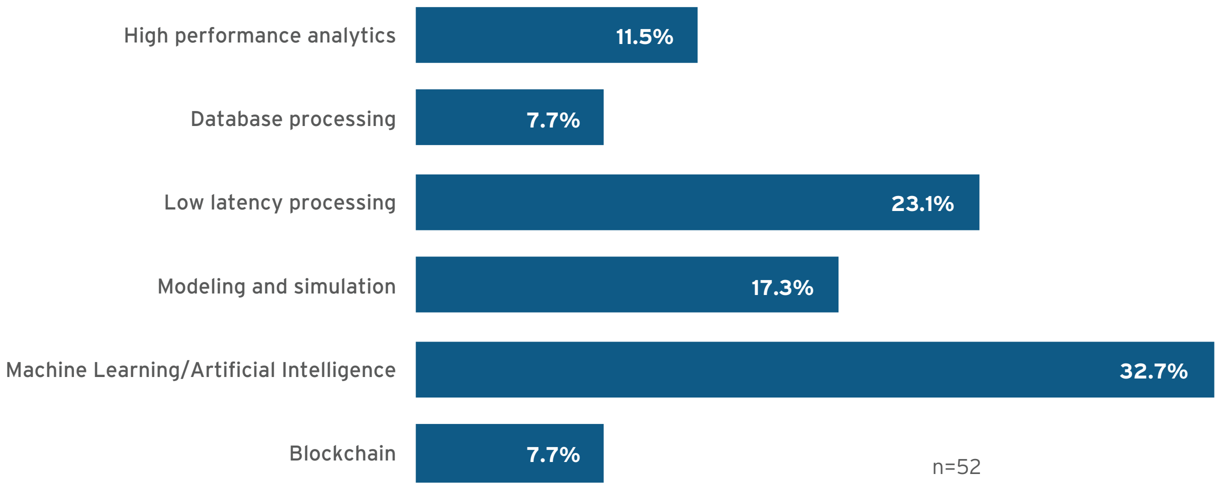
For the reasons discussed above, we see at least three areas that will trigger an increasing number of enterprises and service providers to start adopting DLC. In most cases, we don't initially expect a shift away from air systems, but coexistence where operators use DLC selectively before they consider it for large-scale deployments.

- **Application acceleration**: for a combination of reasons, operators with performance-critical applications will need to look to higher-powered silicon for more hardware performance, particularly if response times are important. Historically, improvements to processor cores have guaranteed double-digit jumps in code performance almost every year, a trend that has largely plateaued as chip designers face fast diminishing returns from further optimizations within a constrained power budget (a few watts per core). IT's saving grace in recent years was the simultaneous uptake of more parallelism in code and flash-powered SSDs that feed previously data-starved processors blazingly fast when compared to HDDs. Software development activity around accelerators (GPUs, field-programmable gate arrays and application-specific chips) has taken off too, but these accelerators already require cooling in excess of 200 watts each, and in many cases, they also rely on a high-performance server processor, which easily brings the cooling requirement past 1kW for a single system.

According to 451 Research's Voice of the Service Provider survey (Figure 1), about one-third of IT service providers said they expect to use hardware acceleration for machine learning and other AI-related applications, while about one-quarter of them have plans for acceleration in low-latency processing (online data mining, live media, financial trading systems and fraud detection, load balancing, etc.). Other areas of interest for more performance are engineering simulations, analytics, databases and blockchain (distributed ledger technology).

Figure 1: Use of GPUs, FPGAs and other accelerators

Source: 451 Research Voice of the Service Provider, Q1



- **Better storage reliability and free cooling**: The leading cause of component failures and application-level disruption is HDDs. Owing to their mechanical nature, HDDs are more prone to fail than electronics at higher temperatures to accelerated wear. Failing HDDs – particularly when there is a systematic issue such as inadequate cooling or a fire-retardant discharge near systems containing HDDs – increase the risk of uncontrolled shutdowns and data losses.

  The corrosive effect of humidity, too, is a major problem because HDDs run at lower temperatures and dissipate less heat than processors, memory, network cards and other microelectronics. In fact, high levels of relative humidity appear to have a much more pronounced effect on hardware reliability than temperature. High humidity levels are likely to cause various issues (condensation and corrosion, and a weakening of the adhesives of silicon packaging, promoting the growth of filaments that result in short circuits, and so on) over time that cause the hard drives' adaptors or controllers to malfunction. Liquid-cooled HDD enclosures and immersion systems can shield HDDs from temperature rises and the moisture content of air, and can also protect them from other external hazards such as vibration from fire agent discharges.

  Using DLC for disk arrays will also set datacenter facilities free from the single biggest (and probably only remaining) technical reason behind tight temperature and humidity controls. In this context, DLC would make air systems not obsolete, but more efficient by supporting the case for wider temperature and humidity ranges. This would allow operators to cut back on mechanical refrigeration – or even eliminate it – and save on electricity and possibly water. This DLC-on-storage approach also offers an alternative strategy to the compute-oriented DLC rollout in environments where there is little need for now to push the envelope of processor performance.

- **Edge installations**: High-performance and more reliable operations across IT and facility infrastructure are what remote edge deployments will require to pack a punch in a small footprint and to minimize visits by field technicians. Using DLC at edge sites – e.g., branch office, retail location, manufacturing floor or a multi-access edge location on a carrier network – should compress the footprint and lower noise and vibration. Due to a lower number of components in the cooling system and more stable environmental controls, the number of failures will be lower compared to an air-cooled system.

  We anticipate that demand at these small localized sites will continue to grow and become more strategic. They are likely to be upgraded from ad hoc installations into dedicated IT installations, with greater compute capacities, yet also higher cooling and power redundancy. They are particularly well suited to self-contained micro-modular datacenters – and for this reason, they can be considered an upgrade opportunity for datacenter technology suppliers. DLC lends itself well to such scenarios, and some DLC rack systems would require a relatively marginal engineering effort to become a standalone edge micro-datacenter.

Conclusions

The technology environment in the datacenter has transformed in major ways over the past five years, and the calculation for using DLC has changed with it. Considerable gains in application performance and lower hardware failure rates will be two key drivers behind the adoption of DLC for mainstream use cases. Ongoing standardization of IT infrastructure and hermetically sealed HDDs will allow operators to phase down air cooling and roll out DLC at scale over time. Vendors of DLC systems have honed their designs for years; they can ramp up production to high volumes and are largely 'IT equipment agnostic.' But much like most technology shifts that have come before, DLC will largely take off because the existing technology – air cooling – will hit its limitations and inflict pain on major buyers of infrastructure in the form of lost performance or higher costs (or both). Datacenters that go operational in 2018 will likely be around in the late 2020s and early 2030s. It is for these reasons that 451 Research views DLC as a must-have in the datacenter operator's toolset if it wants to keep pace with infrastructure requirements.