ToneCurve
Technology, LLC

Moving Beyond the Limitations of Traditional Data Center Cooling

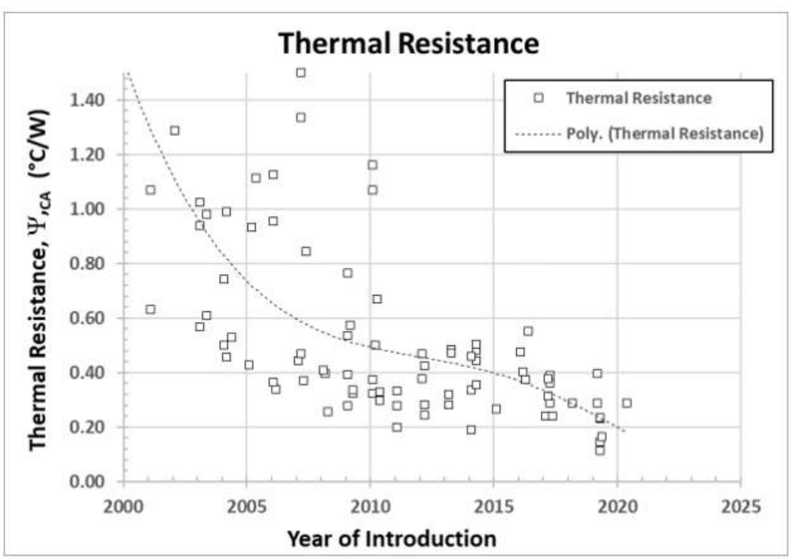
Authored by:
Steven Hill, Independent Technology Analyst

Moving Beyond the Limitations of Traditional Data Center Cooling

By Steven Hill

In spite of the remarkable advancements we've witnessed in the world of computing over the last few decades, one of the most persistent challenges that still remain for datacenter operators lies in the basic problem of removing heat. It's a simple fact that, as servers and other infrastructure devices get faster and more powerful, they also consume an increasing amount of electricity. And as everyone in the IT industry already knows, the majority of the electricity that passes through a computer exits in the form of heat. And for datacenters, it also requires nearly the same amount of energy to remove it.

That's the delicious irony of computing technology; no matter how powerful our systems get one of the key challenges still comes down to the relatively primitive issue of heat management. Of course, that has evolved somewhat over time, especially in the context of massive, hyper-scale cloud datacenters hosting hundreds or thousands of racks of relatively similar and generic servers. But the next generation of datacenters will be increasingly tasked with hosting a growing array of extremely high-performance systems with a dense combination of CPUs and GPUs, designs that are starting to exceed the technological capabilities of traditional methods of heat management in the datacenter (Figure 1 shows the growing challenges with new generations of chips.) Unfortunately, the solutions to the heat problem haven't evolved to the same degree as the demands of high-performance hardware.



Thermal resistance required to cool socket power.

Source: ASHRAE White Paper "Emergence of Liquid Cooling in Mainstream Data Centers"

The Options

Unless you have the luxury of constructing your datacenter facilities in areas of the world where the outdoor temperature remains relatively low all year long, heat management will typically be a key issue. The options haven't changed much over the years until now. Historically, a limited number of techniques have moved in and out of fashion based on the common considerations of cost, efficiency, reliability, and heat load requirements; yet limited to the cooling technology available at the time. Added to this challenge today is a growing concern for sustainability and ecological responsibility.

Air Cooling – Ubiquitous, but Inefficient

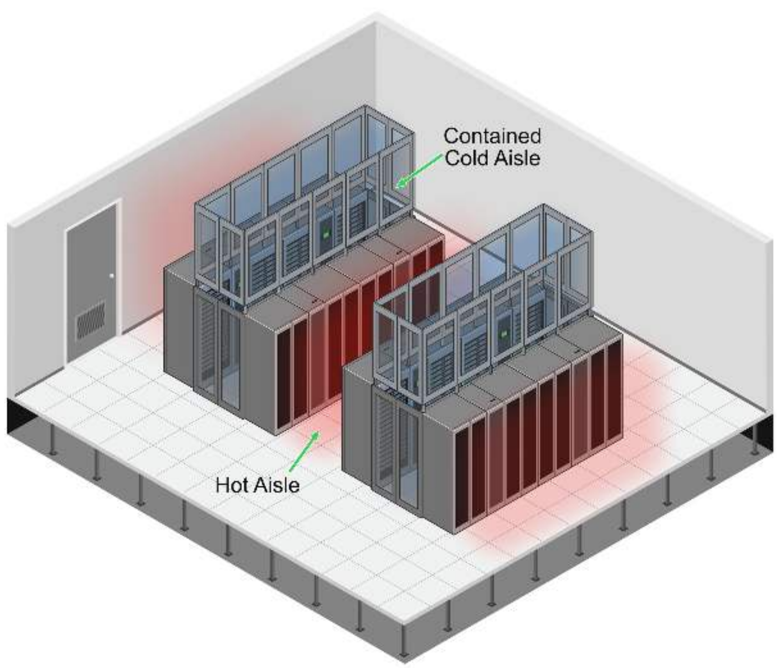
Over the last few decades, the majority of datacenters have continued to depend on the tried and true, raised-floor model of air cooling. Much like a refrigerator, raised-floor cooling depends on the circulation of cooled air throughout an entire machine room. Hot air rising from the systems is gathered from the ceiling and passed through Computer Room Air Handlers (CRAH) or Computer Room Air Conditioners (CRAC); devices that are located at strategic locations at the periphery of the room. CRAHs based on chilled water, or CRACs based on refrigerants cool the air and then deliver it below a tiled floor that's raised with a metal framework between 24" and 48" above a concrete subfloor.

Most raised, or access floors are made up of 24" square tiles that are either blank or have holes or slots cut into them to allow cool air to blow upwards in front of each of the racks. In addition, many raised-floor datacenters also route cabling and plumbing through the same space between the solid floor and the raised floor, which can be seen as both a convenience and a challenge as the layout of the machine room evolves over time. Ultimately, the raised floor concept was designed for a production environment where power consumption was usually less than 10kW per rack, the main problem being that you can only push so much air through the holes in a given tile.

This became a problem in legacy datacenters as the makeup of the racks evolved through the increased adoption of smaller servers configured in higher density within the racks; a design that could easily exceed 20kW/rack. This became more complex because changing the layout of the ventilating tiles can have a cumulative effect on the cooling efficiency of the entire system, which can cause "hot spots" where the local temperature can exceed design parameters. This can be very difficult to detect and remedy and can lead to the need for a complex air-flow analysis study to identify and mitigate problems.

There have been a number of variations on the theme of air-cooled datacenters to address the density problem, including such technologies as fan-assisted floor tiles, hot aisle containment, and auxiliary in-row cooling systems. There are also some highly optimized datacenter designs that combine patented air handling and cooling technologies with concrete slab flooring and overhead cabling. These variations resolve some of the problems due to primary air cooling, but air cooling alone remains somewhat limited because air is simply an extremely poor conductor of heat. In fact, air is a main component of many insulation technologies; so simply cooling the air and blowing it at a heat source may actually be the least efficient model for cooling there is.

Of course, the majority of high-density datacenter operators have adopted air as the model for cooling massive, cloud datacenters; likely because the substantial evaporative cooling solutions they require continue to be energy-efficient at scale for moderate work loads. But this also presents problems because of the extraordinary water consumption requirements of these facilities; a large cloud datacenter can consume billions of gallons of fresh, potable water per year. This is becoming problematic in drier parts of the world that are attractive to operators because of the relatively low cost of real estate, but also where water availability may already be a serious concern. Ultimately, hosting an increasingly hot and dense infrastructure in a large, air-conditioned room is unlikely to be sustainable forever.



Example of Hot / Cold Aisle Containment

Water -based Direct Liquid Cooling – Better, but Challenging

Water-based, or single-phase, cooling is the grandfather when it comes to cooling large-scale computing platforms, and was a requirement for a number of large mainframes from the 1960's and 1970's. Water is about 24x more efficient than air when it comes to transferring heat, which makes it a natural for a number of cooling applications. Today, direct to chip, water-based cooling is an option for consumer and enterprise-based cooling applications, and as if to highlight the point, IBM introduced the IBM z15 in 2019; its first water-cooled mainframe design in over two decades.

IBM was not alone in the return to H2O, there are a number of OEM and ODM vendors that now have water-cooled server offerings, as well as aftermarket add-on kits based on a closed-loop design with an external radiator. But water-based cooling isn't without its challenges, simply because water and electricity really don't mix particularly well. I know this from personal experience, as the person who was directly responsible for recovering from a catastrophic datacenter flood in 2005. The Network Computing Real World Test Lab where I operated as Servers and Storage Editor was drenched due to the failure of a water heater in the business directly above; and it took out nearly 8 racks of servers, networking, and storage hardware. It took almost a year to recover...a year in which I learned more about the intersection of water and electronics than I ever wanted to know. (Pro tip: It's not as much about the water itself, rather the impurities dissolved in it that wrecks electronics.)

The majority of in-server water cooling is based on the "direct-to-chip" (D2C) model, where pressurized water is passed through a water block covering the chip and then

returned to a radiator or a chilled water source for recirculation. Those utilizing a radiator/fan combination are typically closed-loop designs, where the plumbing is pre-integrated, factory sealed, and requiring no maintenance by the end user. There are also a number of enterprise-level systems are instead based on purpose-built server designs that incorporate integrated plumbing, customized heat sinks, and a system of rack-level water distribution and recovery manifolds that connect directly to an external water chiller system.

D2C water cooling is clearly an improvement on air-only cooling for hosting higher-density applications within an existing air-cooled datacenter, but it's hardly common and almost never used as the primary source of cooling. One of the few exceptions I've experienced is from OVH (OVH Groupe SA), one of the largest cloud providers on earth who have integrated D2C as part of their self-manufactured, proprietary servers. These systems are based on customized chassis, racking, CPU coolers, and a datacenter-wide water distribution and recovery system.

While somewhat unique in the typical datacenter, D2C water cooling not only benefits from the greater efficiency of water for transferring heat, but also because D2C can specifically target the CPU and GPU, often the largest point-source of heat in the system. However, the idea of allowing water anywhere near their powered infrastructure is anathema for a large number of administrators, and water-based cooling requires a management model that's capable of assuring temperature control and flow monitoring, as well as water treatment and drains or catch basins to mitigate potential leaks.

Adopting water for extremely high-density cooling also has some limitations, in that water has a finite ability to cool the substantial amount of heat generated over the small surface area of the CPU/GPU. In that water can at most be chilled to just above freezing, the only way to increase water's cooling potential is to increase its flow rate substantially. Testing has shown that it would require a flow rate of ~2.2 litres/minute to cool a 500 watt processor, which would require a substantial amount of pressure forced through thin tubing when compared to other liquid-based cooling options. Water-based solutions also require additional maintenance & service to prevent biofouling and contamination which can lead to system failures and downtime.

Dielectric Liquid Cooling – Evolving as a Viable Option

One of the earliest uses for refrigerant cooling was the Cray 2 supercomputer that was introduced in 1985. An incredibly powerful system for its time, but the only way to keep the extremely high-density circuit boards cool was to bathe them in a flowing dielectric coolant that wouldn't react with energized components. The Flourinert coolant developed by 3M was pumped through the chassis and then bubbled up through a dramatic, see-through "waterfall" chamber, almost as large as the computer itself, where the refrigerant reverted back from gas to liquid for recirculation.

The Cray 2 Supercomputer with cooling tower at the Computer History Museum – Source: Wikipedia

It seems that everything old is new again on a revolving basis in IT, and the same holds true for refrigerant-based cooling. Based on a similar principle as our air conditioners and refrigerator-freezers this model leverages a cycle where heat is extremely efficiently transferred through the evaporation and condensation of a fluid with a relatively low boiling point. Today there are several models evolving for liquid cooling.

<u>Immersion Cooling</u>

The modern version remains very similar in concept to the Cray design, in that entire circuit boards are specially designed to be completely immersed in dielectric coolant. There are two forms, and both require a substantial tank filled with liquid, with the main difference being whether or not it's based on the evaporation/condensation process of refrigeration.

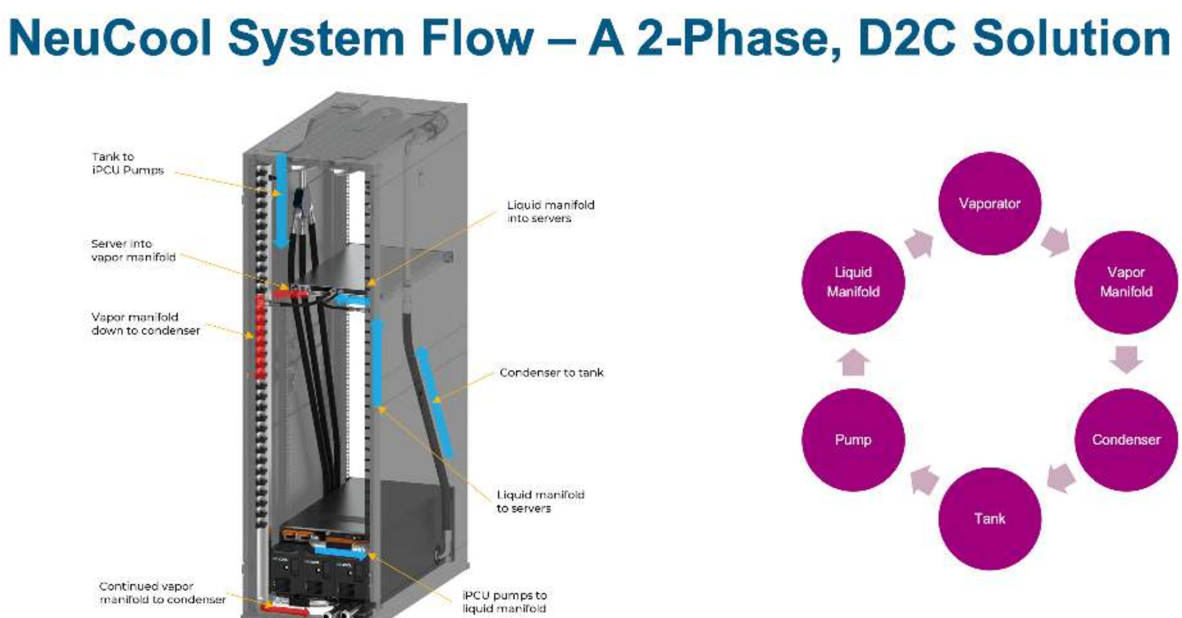
- **Single Phase** – A single phase system is based on a non-refrigerant dielectric fluid that circulates through multiple immersed server boards and pumped through a conventional heat exchanger connected to a chiller system to maintain the desired overall temperature. The benefits of this system come from both the greater heat transfer capabilities of liquid, as well as the ability to capture the majority of the heat generated by the server. There are multiple options for oil-based dielectric liquids that are not refrigerants. The main drawback to the single-phase system is that, while much better than air at removing heat, it can cool about half the wattage of a two-phase immersion alternative.
- **Two-Phase** – This is the most visually striking model, where you can watch the fluid-covered circuit boards bubbling away as the refrigerant changes from liquid to gas. Then there is a cooled evaporative loop located in the same tank as the components, where the heated gas condenses on cooling elements within the tank and returns to cool again. In a two-phase system, there is the additional benefit of an even higher cooling yield caused by leveraging the phase change cycle.

While both of these models benefit from the direct contact of fluid with the components, they also present some challenges. They typically require a large tank that recirculates hundreds of gallons of cooling liquid which could potentially weigh a ton or more and in the case of two-phase dielectric fluid, can cost upwards of $100K to fill the tank. In addition, it requires all components to be certified immersion-ready, including data and power distribution cables. It also requires specialized lifting equipment for servicing components; not to mention a system for transferring, filtering and coolant disposal and spill remediation.

<u>Refrigerant-Based Two-Phase D2C Cooling</u>

This is the latest entry in the challenge for high-density datacenter cooling and offers some of the best attributes of earlier options; in particular the way it leverages the heat transfer enabled by the phase change of the refrigerant fluid from liquid to gas and back again in a closed cycle. Like a water-based D2C system, primary cooling is provided by a cold-plate heat sink that is attached to the primary heat sources such as CPUs, GPUs, and potentially any other high-density ICs that may become major heat generators. However, unlike a water-based system, two-phase can remove heat by taking advantage of both sensible & latent heat of vaporization. It also requires 4X-9X less flow rate at the chip compared to single-phase, water-based cooling.

Accelsius' NeuCool Platform is an emerging example of this technology. Their process is specifically optimized to remove heat from relatively small electronic components. The cold-plate—also known as a "vaporator"—in contact with the processor is supplied with a liquid refrigerant, where it converts to gas in the presence of heat. The warmer gas flows through a manifold as part of a closed loop to an integrated heat exchanger where it condenses by transferring its heat to facility water or another refrigerant. The cool liquid refrigerant is pumped through a second manifold back to the vaporators to continue the cycle. An attached reservoir allows for expansion and contraction of the refrigerant under varying compute power levels, and provides adequate suction pressure for the pumps. NeuCool uses an intelligent Platform Control Unit (iPCU) to house the pumps, heat exchanger, monitoring & control system, sensors, and most electrical and electronic components associated with the system. This is similar to a coolant distribution system (CDU) used in other D2C systems.



Source: Accelsius

This type of cooling is ideally suited for high-density, rack-level integration; where the CDU, heat exchanger, and all required plumbing are designed to operate as a stand-alone system. However, this design can also be used for partial racks, or extended to multiple racks in a scale-up or scale-out model by simply increasing the capacity of the CDU supplying the system, and without introducing any major impact to the rest of the datacenter environment. The ability of refrigerant-based two-stage D2C cooling to offset about 80% of the heat load makes it a viable technology for extending the life of existing air-cooled datacenters with the need for new HPC or AI infrastructure, or for making it possible to host high-performance server clusters non-traditional datacenter environments such as mission-critical edge deployments.

Even though refrigerant-based two-phase D2C cooling technology could be used in almost any context, the sweet spot for such effective cooling is likely best focused on the highest-performance applications, such as high-speed trading, real-time analytics or to accommodate the growing use of AI that's leveraging some of the largest computing environments today. For example, single Nvidia DGX H100 supercomputer fills 6-RU of rack space and maxes out at just over 10kW—not to mention a host of multi-node HPC designs that can draw ~3kW in only 2RU—and it's becoming more and more common to see rack designs that can exceed 50kW.

The Conclusions

Like many things in life, there isn't really a one-size-fits-all option for datacenter cooling. So much depends on the complex aspects involved, such as: budget, natural resources, infrastructure choices, workload requirements, energy costs, support issues...the list can go on and on. While traditional, evaporative air cooling remains the most common form of heat management for large-scale environments, the challenges posed by the increased density and power consumption of next-generation infrastructure can have a major impact on datacenters of every size.

The fact is, what works at cloud scale may not be at all functional for smaller and more diverse datacenter environments. In conversations with administrators from a broad range of environments, the needs of medical, business, educational, scientific, engineering, financial, and communications disciplines can vary dramatically, even between those in a similar environment. And in some cases, needs can change at any time, and unfortunately, not every technological decision is in the control of the IT group; which can force IT groups to scramble to accommodate new requirements by any means necessary. It has historically been very difficult to shoehorn a high-performance cluster into an existing facility, and until now the options have been somewhat limited.

When looking for a solution for any IT problem I rely on a simple but effective checklist:

- [ ] Is it resilient?
- [ ] Is it compatible with the existing physical environment?
- [ ] Is it easy to maintain?
- [ ] Does it work with our existing management tools?
- [ ] Is it compatible with a variety of vendors?
- [ ] Is it safe?
- [ ] Is it scalable (both up and out)?
- [ ] Is it cost-effective?

If you can answer yes to the majority of those questions, then I believe you're in pretty good shape. It's good to have options, and it's unfortunate that as technologists, we don't always have very many. The availability of a highly efficient D2C cooling system seems to be tailor-made for a rapidly-evolving computing environment; and I don't see the trend of increased density and power consumption going anywhere but up. I believe that, for the forseeble future, we're facing a veritable infinite loop of information, results, and more information. And heat. A lot of heat.

Steven Hill is an Independent Technology Analyst and the owner of ToneCurve Technology, LLC.

For over two decades he written hundreds of articles, analyst reports and spoken on evolving data center technologies. In his various positions as both a journalist and an analyst he has focused on a wide range of topics, including: servers, storage, converged infrastructure, data protection, disaster recovery, business continuity, power management and cooling technologies. The opinions expressed in this paper are entirely his own, and may not reflect the position of our sponsor.

This paper has been sponsored by Accelsius, you can find more information on the NeuCool system and how we check all the boxes outlined above. Accelsius.com