Let the water flow!

28 November 2016 **By Ben Coughlin, Colovore**

Despite its clear efficiencies in cooling capabilities versus traditional air-based systems, liquid cooling adoption in the data center has been slower than anticipated. But this is changing, and it is changing quickly.

With computing infrastructures continuing to miniaturize while processing capacities increase, power requirements per rack unit of IT infrastructure are growing significantly. With the advent of hyperscale computing, HPC architectures, Big Data and Hadoop clusters, and converged infrastructure deployments it is becoming increasingly common today to see 1U servers drawing 500W to 1kW of power, an increase in power draw of 20x since 2000.

This presents challenges inside data centers, especially in multi-tenant colocation facilities with a wide variety of IT infrastructure requirements to be satisfied. Higher performance infrastructure generates considerably more heat per rack unit and per data center cabinet… so what is the best way to cool it?

Liquid.

Cooling advantages versus air

Colovore has a 9MW liquid cooled data center in Santa Clara

Source: Colovore

Large web companies and data center operators are increasingly deploying water-cooled systems, because liquid has a big advantage over air. Water or coolant has 3,000x more capacity to remove heat than air, driven by its superior heat capacity, thermal conductivity, and density. To put it simply, water can absorb more heat than air and it does it faster.

The practical advantage of liquid versus air is that when deployed the cooling medium ends up much closer to the source of the heat generation (ie, the server). Hot air molecules can't disperse and escape.

In summary, then, we have a more effective cooling medium being deployed in closer proximity to the source of heat. Voila! Major efficiencies gained.

Liquid cooling methods

To date, the majority of liquid cooling systems being deployed involve the use of heat exchangers, which are coils of water that cool hot air blowing out of servers. IBM developed this technology years ago, and these systems today either exist on the rear door of the cabinet or in some cases above the cabinets. There are numerous examples of heat exchanger-driven data center deployments ranging from hyperscale deployments (Google uses ceilings of water coils to cool its contained hot aisles) to large web providers (eBay in Phoenix and Linkedin in Oregon, which both use rear doors) to modern colocation facilities (Colovore in Silicon Valley, which also uses rear doors).

Other forms of liquid-based cooling systems remain in earlier stages of deployment, but there has been progress. In a targeted water system, water is passed through micro-channels, and CPUs and GPUs are cooled by liquid apparatus inside the server. In liquid immersion systems, rack-mounted servers are installed in special tanks filled with coolant, and the liquid "bath" absorbs and circulates the heat to a heat exchanger, where it is extracted. Recently [Dell announced its Project Triton system for eBay](http://www.datacenterdynamics.com/content-tracks/power-cooling/dell-reveals-project-triton-water-cooling-tech-in-use-by-ebay/96312.article) (URL=http://www.datacenterdynamics.com/content-tracks/power-cooling/dell-reveals-project-triton-water-cooling-tech-in-use-by-ebay/96312.article) , which is a custom-built solution featuring water delivered directly to a cold plate on top of CPUs in the server chassis. The [University of Birmingham in partnership with Lenovo](http://www.datacenterdynamics.com/content-tracks/power-cooling/lenovo-introduces-water-cooling-at-university-of-birmingham/96270.article) (URL=http://www.datacenterdynamics.com/content-tracks/power-cooling/lenovo-introduces-water-cooling-at-university-of-birmingham/96270.article) has similarly announced a liquid cooling system in which water is delivered inside the server to direct-attached heat sinks on the CPUs and RAM.

Why now?

Historically there has been resistance in deploying liquid cooling because in general water and electricity mixing together is not a formula for success, and it is very difficult to retrofit legacy data center environments to move to a water-based system given the piping (ie, water distribution) required to be built and integrated.

But this is changing, as we have seen from the previously-mentioned examples. Why now?

First, advances in server processing capabilities—in other words the advent of higher performance computing infrastructures—are driving power densities per rack higher. At some point, air-based systems cannot physically cool the heat being generated by these systems, and water is the most efficient and effective manner to deliver higher densities. At our facility, for example, we routinely see our customers require up to 20kW of critical power per rack. For HPC, deep learning, artificial intelligence, and Big Data architecture deployments this is an important consideration.

Second, water-based systems provide a very flexible manner to accommodate variable customer power requirements. The same water distribution loop, cooling tower, pumps and rear doors can deliver cooling capacities to support 1kW of load and up to 35kW of load– it's simply a function of water loop temperature and flow rate (ie, psi) to each rack. This is in stark contrast to air-based systems in which cooling capacities are mostly fixed per square foot. Water provides colocation operators much greater flexibility in accommodating diverse requirements, in particular supporting higher power density requirements. The "rob Peter to pay Paul" phenomenon in which a high-density section of an air-cooled data center floor can be provisioned but only by stealing cooling capacity from another section of the floor and rendering it unusable is no longer an operational limitation.

> Pumps pushing low-pressure water through pipes are more reliable and easier to maintain than a network of CRACs

This also provides a major advantage as far as customer scalability goes—scaling power and cooling within the same rack by water temperature and flow rate versus needing more physical space is a significant advantage for a growing customer. In tight data center markets like Silicon Valley, for example, if growing compute requires more and more floor space eventually there is no more space to grow into and a customer gets stuck. IT executives can sleep better at night when thinking about how they will grow inside a liquid-cooled data center.

Third, water-based systems are simple to operate. Fundamentally there are pipes and pumps and that's the bulk of the cooling system. There are very few moving parts and little annoying things that tend to break down as with CRACs (read: belts, fans, compressors, pulleys). Pumps pushing low-pressure water through pipes are more reliable and easier to maintain than a network of CRACs with all kinds of components to manage.

Fourth, because of the inherent nature in cooling efficiency, water-based systems significantly improve power usage effectiveness (PUE). Facebook's 1.06 PUE. in Prineville is wonderful but we have to recognize that Facebook has deployed server and data center architecture that is entirely customized for its specific purposes. In a multi-tenant data center environment, not all customers are Facebook. But, with water-based cooling, a PUE. of 1.1 to 1.2 is readily achievable, representing significant cost savings for the operator and quite beneficial from an environmental and energy perspective versus traditional air-based systems.

Ben Coughlin is chairman and co-founder at [Colovore](http://www.colovore.com/) (URL=http://www.colovore.com/)