Liquid cooling applications

A Flexential guide

The purpose of this guide is to provide an overview of how Flexential liquid-ready data centers can support liquid cooling applications today as machine learning workloads and other GPU applications are becoming the predominant users of data centers. Today, **Flexential Data Centers** support air-cooled workloads of up to 50 kW per cabinet. Using the liquid cooling techniques described in this paper, high-density cabinets can be supported by liquid cooling to the server, and high-power use servers can be supported via immersion to support workloads of up to **80 + kW per cabinet.**

Disclaimer: Throughout this paper and in general conversations regarding data center operation, the term cooling is commonly used to describe the process of keeping IT equipment operational and not overheating. In purely technical terms, there is no such thing as cooling. There is only the absence of heat. The technical description of the process used is heat rejection, where heat is removed from the servers (via air or water) and conveyed to the exterior of the facility and rejected to the atmosphere, and the process is repeated. For ease of use, the colloquial term cooling is synonymous with the technically accurate term heat rejection.

To learn more about how data center cooling improves performance and avoids downtime and unforeseen issues, read our blog.

History

Liquid cooling entered the computer sector early in its history when IBM released a series of enterprise-grade computers called System/360 in the early 1960s. Mainframe computers developed by Control Data Corporation (CDC), such as the Cyber 170, were also water-cooled.

Liquids, specifically water, are significantly more efficient at transferring heat than air. Water has 24 times the heat-transfer capability than air, and various forms of liquid cooling have been in use for years. Water is supplied from a central cooling plant via pipes installed under the raised floor. The central processing units of the mainframe are connected to this piping, and the heat generated by the processor is transferred to the facility's water piping system.

Liquid cooling first entered the computer industry in the early 1960s with IBM's release of the enterprise-grade System/360 computers.

Consumer-grade liquid cooling

Today, liquid cooling is present even in a typical desktop PC—and the concept has essentially remained the same. **The cooling process is made up of three distinct parts: the heat plate, the supply and return pipes, and the radiators and fans.** The heat plate is essentially a metal plate that covers the CPU die with a small reservoir on top. The plate is engineered to be as conductive as possible in terms of heat. Any heat generated by the chip will be transferred to the reservoir on top.

The liquid in this closed loop will travel via the supply and return pipes to the radiators, where heat will be pushed out of the PC enclosure through the radiator fins—these fins are actively cooled by fans. Consumer-grade liquid cooling options have originally only dealt with CPU heat, but **now almost every component of a modern-day PC can be liquid-cooled.**

Liquid cooling equipment

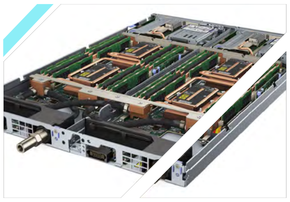
There are **three main methods of implementing liquid cooling** in data centers. The first to be used at scale was a rear door heat exchanger. The second method is immersion, and the third is direct-to-chip, sometimes referred to as LCDC.

Rear door heat exchanger

In a **rear door heat exchanger (RDHX)** application, air-cooled servers are installed in a cabinet, and a heat exchanger is installed on the rear of the cabinet. The heat exchanger is comprised of a set of radiator fins that are connected to a cold or chilled water supply. A set of fans pulls air from the cabinet over the heat exchanger, and the heat from the servers is transferred to the liquid in the heat exchanger. This heat is then rejected by the facility's water system. RDHX are typically powered with generator-backed utility power but can also be powered by critical UPS power. Typical RDHX are not redundant, and in some applications, a failure of the RDHX can impact the servers in the cabinet.

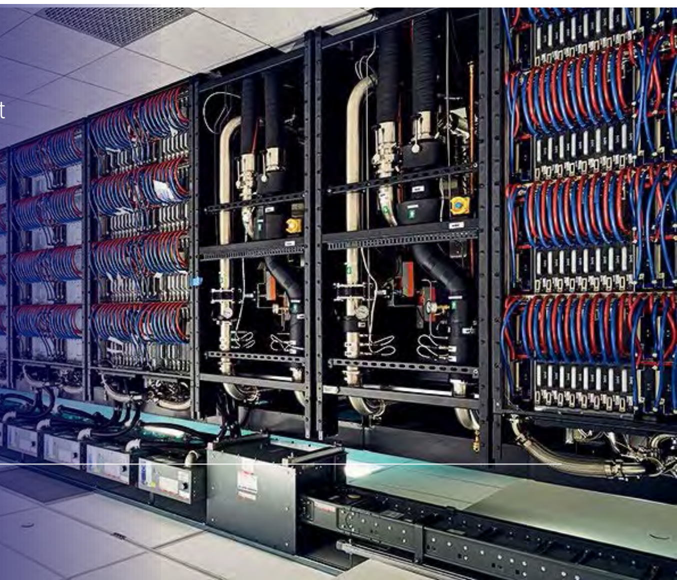
Immersion

As the name implies, **immersion cooling** is a rejection technique where a traditional not-liquid-cooled server is immersed in a tank that is filled with a dielectric fluid. The fluid in the tank then serves as the heat-transfer mechanism by removing heat from the server. The fluid in the tank is then circulated through an internal heat exchanger, where the heat from the tank is transferred to a facility water loop. The benefit of using immersion is that the servers do not have to be modified or specifically designed for liquid cooling. The installation and use of immersion cooling requires a gantry to be constructed over the tanks to provide a mechanism for the servers to be hoisted out of the tank and then left in the vertical position to allow the tank dielectric fluid to drain off—a "drip dry" process. The primary benefit of immersion is that 100 percent of the heat created by the server is captured into the facility's water-cooling system.



Direct-to-chip

Direct-to-chip, also known as cold-plate cooling, is an approach that attaches small radiators, called cold-plates, to the major heat-producing elements of a server, typically CPU, GPU, and memory.
The cold plate has small capillary lines that are connected to a manifold in the server, and the manifold supply and return lines are routed to connectors on the rear of the server.



The key elements of a liquid-cooled computer include a Cooling Distribution Unit or CDU, represented by the two cabinets on the right side of the image. The computational elements are then connected via larger stainless piping sections and individual liquid connections to each element via the red and blue hoses.

Pictured: Aurora supercomputer, Argonne National Laboratory

Liquid cooling advantages

The following are the **four primary advantages** of using liquid cooling for advanced IT workloads.

Reliability and performance

Air cooling simply can't keep up with the growing density and heavy processing loads. With liquid cooling, you don't need to throttle back CPU and GPU performance as much to prevent thermal runaway. **Greater cooling power means greater performance** from your HPC hardware.

Efficiency

In addition to being dramatically more efficient at transferring heat than air, liquid cooling allows for far more precise temperature control with in-rack CDU and on-proc cold plate technologies. Instead of running coolant to the entire system all the time, you can **run it directly to the racks that need it,** reducing operating costs.

Density

With air-cooling, you limit the density of your data center. The increased cooling power of liquid cooling allows you to **maximize the use of your data center space,** achieving greater processing potential in smaller spaces and avoiding the need for costly remodels or expansions.

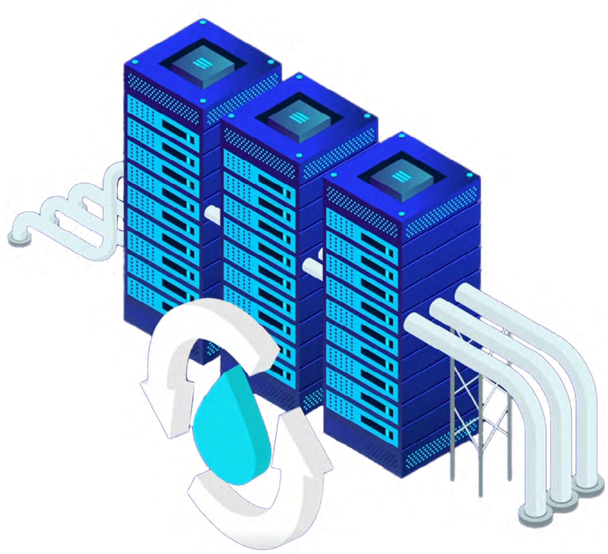
Sustainability

Each year, data centers consume approximately two percent of global power consumption. Liquid cooling not only improves sustainability by helping data centers reduce energy consumption but also allows them to **repurpose captured heat more effectively.**

Flexential liquid-ready data centers

Flexential has been constructing data centers that support liquid cooling since 2015. **Our data centers support liquid cooling** via the use of a **heat-rejection module or HRM, liquid cooling via air-cooled chillers, and facility-wide chilled water.** Rear door heat exchangers, CDU, or immersion tanks can be directly connected to the facility water loop.

Whichever solution you tap into, special arrangements are made to ensure the integrity of the secondary cooling loop. In the case of direct-to-chip, the pumps for the CDU and primary loop are powered by dedicated UPS or by power from the critical IT UPS plant. This ensures that both the primary and the secondary cooling loops have reliable flow for the server loads.



Liquid cooling to ensure high performance and low latency

Liquid cooling solutions for next-gen workloads

The evolution of liquid cooling applications has marked a significant milestone in the data center industry, offering **innovative solutions to address the rising demands of advanced workloads.** As highlighted in this guide, our commitment to being liquid-ready positions us at the forefront of this transformative technology.

From the historical roots of liquid cooling in mainframe computers to the present-day consumer-grade liquid cooling in desktop PCs, the journey has been characterized by increased efficiency and precision. Today, **Flexential data centers lead the way in supporting liquid cooling for high-density cabinets,** allowing for the effective management of power-hungry servers through techniques like immersion and direct-to-chip cooling.

The advantages of liquid cooling, including enhanced reliability and performance, greater efficiency, improved density, and a positive impact on sustainability, make it a compelling choice for data centers looking to optimize their operations.

Let Flexential help you create a more efficient data center environment. Reach out to our team for assistance with liquid cooling solutions. We are committed to supporting you and answering any questions you may have.

Schedule a Consultation

As we continue to embrace the future of data center cooling, Flexential remains dedicated to staying at the forefront of innovation, meeting the evolving needs of our customers, and contributing to the sustainability of the industry.

Download our latest ESG Report to learn how we are maximizing efficiency, minimizing our carbon footprint and water usage, and continually investing in updates and improvements that strengthen our sustainability efforts across our fleet of data centers.