

DIRECT LIQUID COOLING AND ITS IMPACT ON THE DESIGN OF MISSION CRITICAL FACILITIES: An Introduction to Hybrid Dry Adiabatic Cooling™

R. Stephen Spinazzola, PE, LEED AP

stephen@shumateengineering.com

*Shumate Engineering, PLLC
1577 Spring Hill Road, Suite 510
Tysons, Virginia 22182*

This paper explores the evolution of data center cooling technology to address the anticipated demand for increased rack density for IT equipment in the white space. In our experience, most of the hyperscale data center operators have reached a density limit for IT equipment, due to the inherent limitations of air-cooled cabinets. On average, much of the industry is achieving rack density as high as 11 kW/cabinet with current designs. The emergence of artificial intelligence processing requirements and further density increases for cloud computing will continue to push that limit. We anticipate that many data centers will require significantly higher density moving forward, with some systems requiring as much as 40 kW/cabinet in the near future.

Cooling IT equipment with direct connection to chilled water has existed since the mid-1960s, dating back to some of the original water-cooled products from IBM. Traditionally, the mainframes used chilled water and a heat exchanger to cool the processors. The remainder of the equipment was air cooled using chilled water Computer Room Air Handling Units (CRAHUs). That approach was used throughout the 1980's and 1990's for large supercomputers. However, due to concerns about water in data centers, direct connections to chilled water systems were largely eliminated from commercial and enterprise data center designs over the last 20 years. The current industry is heavily reliant on air-cooled racks in mission critical applications.

As we advance the clock to 2023, various technologies for liquid cooling are becoming mainstream. They come in a wide range of deployments and applications, including liquid cooling of single server, liquid cooling an entire rack, or immersion cooling with servers submerged in a non-conductive fluid. These liquid-cooled technologies can collectively be referred to as Direct Liquid Cooling (DLC).

For any data center design, the gold standard used to evaluate cooling technologies is Power Usage Effectiveness (PUE). PUE was adopted by Green Grid on February 2, 2010. Green Grid is a non-profit organization of IT professionals. PUE has become the most commonly used metric for reporting the energy efficiency of data centers¹. This factor represents the Total Facility Energy Use divided by the IT Equipment Energy Use: the closer to 1.0, the better the PUE. This paper will utilize PUE when discussing relative efficiency and performance of various mechanical technologies.

Current Design Standards – Challenges and Limitations:

In today's market, most high-density mission critical projects are designed with direct adiabatic cooling, using large fan arrays to distribute airflow throughout the white space and condition the IT rack equipment. In this scenario, equipment is typically configured using N+1 redundancy criteria, where (N) is the required cooling units plus one more in various configurations. This approach is more efficient than a once prevalent 2N redundancy, where the total required capacity is doubled. For comparative

purposes, we have calculated the performance of this example system to include several typical design features. These include hot aisle containment structures, increases in ASHRAE cold aisle temperatures as high as 80.6° F, and the use of direct adiabatic cooling AHUs. These design options are very common in modern data center design and are essentially standard practice for most hyperscale data center clients. There are relatively few remaining options to significantly lower the PUE of new all-air cooled data centers.

Speaking with industry leaders in the design and operation of data centers, the anecdotal consensus is that direct adiabatic cooling cannot effectively cool densities beyond 11 kW/cabinet. This is largely attributed to the target air temperatures and the inherent airflow limitations within a traditional rack cabinet. In addition to the capacity concerns, the sheer volume of air circulation inside the data hall introduces more possibilities of contamination that can impact IT reliability.

For example, the airflow required to condition 1 kW of IT load (assuming air delta T = 20° F) is approximately 158 cubic feet per minute (CFM). While not an excessive number on its face, an 11-kW rack would require 1,738 CFM to condition the IT load. This translates to 2.5 million cubic feet per day for a single rack. This not only requires a tremendous amount of fan energy, but when airside economy functions are introduced to conserve energy, there is more risk of contaminants and particulate in the airstream.

For hyperscale data centers with a constrained site, the roof becomes a huge array of mechanical equipment. Recently Shumate Engineering was hired as a Subject Matter Expert to help review a new hyperscale data center project. The proposed mechanical system included 30 direct adiabatic AHUs on the roof, with over 2,500 lineal feet (nearly one-half mile) of roof curb directly above the white space. The number of roof penetrations creates dozens of additional roof-leak points of failure for the system.

Thus, we conclude that the current designs for air-cooled IT racks will not be a viable solution to meet the density requirements in the future. As both the Engineer of Record and Subject Matter Expert on several high-density data centers, we have both empirical and anecdotal data that all-air solutions have hit the proverbial brick wall.

We have calculated the energy use for direct adiabatic cooling for a variety of IT loads, ranging from 10 to 25 MVA. The energy used for the supply and return fans at an N+1 level of reliability is approximately 0.092 KW per 1 KW of IT load. This generates a PUE of 1.092 for just the fans. When adding all the other elements that contribute to the total PUE, the calculated Power Usage Effectiveness is more than 1.120.

Direct Liquid Cooling Technology:

Direct liquid cooling is not a perfect fit for all data centers. History tells us that new technologies take time for designers and operators to accept and implement. For example, hot aisle containment was first introduced to the data center world in 2001 at a 7X24Exchange conference by team members from Shumate Engineering. It took roughly five years to gain significant traction in the market before it became a new industry standard.

There are several technologies that are being used to implement Direct Liquid Cooling. Depending on the client and the process requirements, end-users may have a preference for one versus the other.

Immersion Cooling (IC) – The process of Immersion Cooling utilizes liquid filled tanks where the servers are designed and deployed in horizontal racks that are fully submerged in a tank with a non-conductive fluid to absorb the heat. The solution provided by various Immersion Cooling companies includes a liquid-to-liquid heat exchanger, designed to transfer the energy from the submerged servers via the non-conductive fluid to the process water and then to the heat rejection equipment. There are several types

of devices that can reject heat to the atmosphere. Cooling towers are commonly used to dissipate heat but are the highest consumers of domestic water in a cooling plant design. Cooling towers are 100% adiabatic (evaporating water) and usually include a bleed line that is equal to the evaporation rate. This bleed line is used to reduce the build-up of scale (mineral deposits) on the internal components of the cooling plant, but the process doubles the use of domestic water.

Rack Fluid Cooling (RFC) - The other technology available for Direct Liquid Cooling includes fluid piped directly to the server in a conventional rack, which we call Rack Fluid Cooling (RFC). Each server is equipped with a “cold plate” heat sink that transfers the heat from the microprocessor to the fluid. The cooling fluid is circulated through a closed loop with a liquid-to-liquid heat exchanger designed to transfer the energy from the RFC to the cooling fluid closed loop. This system also rejects heat into the atmosphere using a cooling tower or other device, similar to the Immersion Cooling system.

Regardless of the specific technology, data center operators are asking “when do I need to start implementing liquid cooling in my data center?” More importantly, they are wondering how to integrate a liquid cooling solution when so much of their infrastructure is based around an all-air-cooled system. The basic answer depends on when your data center IT load will exceed the average of 11 kW/cabinet.

Shumate Engineering has conducted extensive research on the use of direct liquid cooling for high density mission critical facilities. We are at an inflection point due to two critical components in the design and operation of these data centers:

- The inherent rack density limitations of all-air cooled systems to cool these facilities.
- The high energy usage cost and water usage with one of the most popular designs for an all-air data center - direct adiabatic cooling.

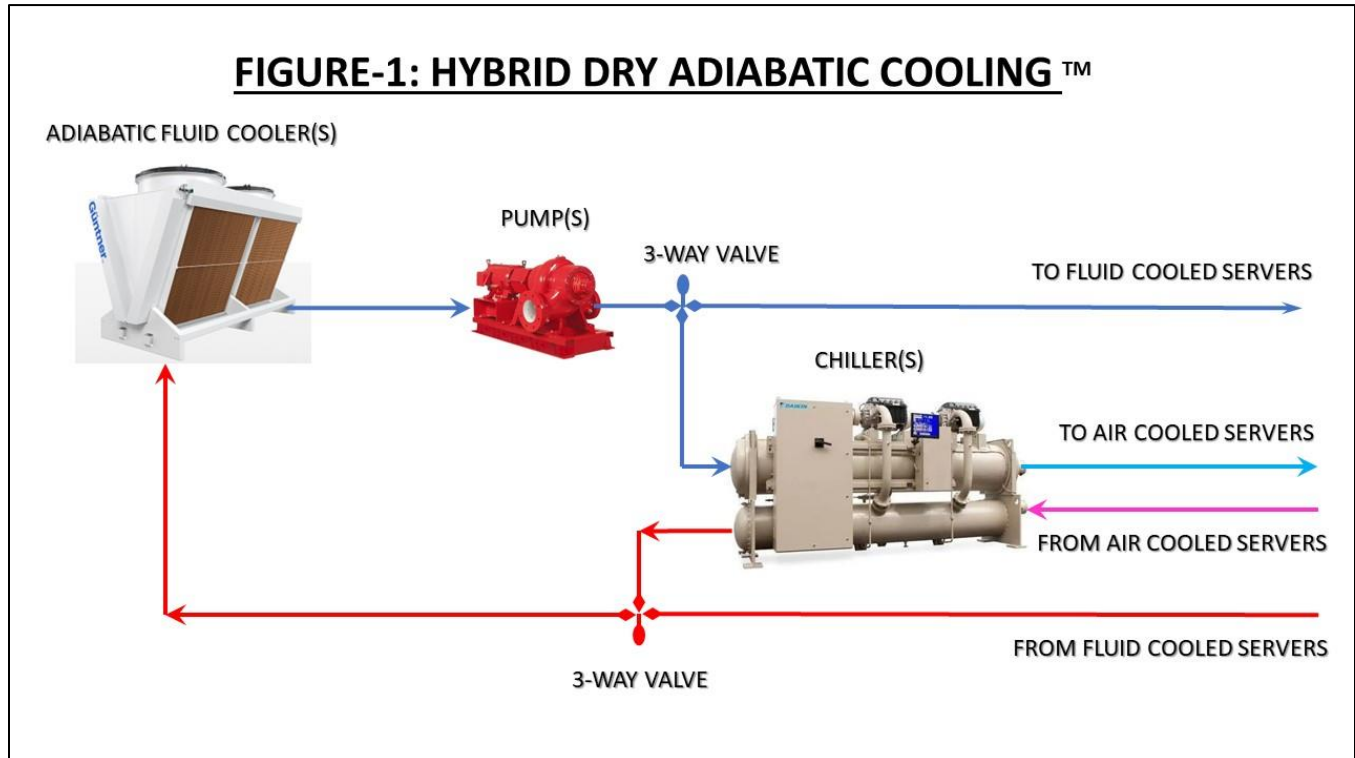
We believe the next generation of data center design is directly tied to the available operating temperatures for Direct Liquid Cooling technologies. In Immersion Cooling, the cooling fluid entering temperature can be as high as 37⁰ C (98.6⁰ F). As recently as November 2022, Hewlett Packard² announced their complete liquid-cooled product, offering a packaged solution direct to chip liquid cooling. This packaged product includes a heat exchanger to the liquid distribution for the racks and liquid distribution to each server. The product is operable with inlet water temperatures of up to 32⁰ C (89.9⁰ F).

The ability for these systems to condition high density IT load is the primary benefit. However, from an operational perspective and when calculating the total PUE, the most intriguing aspect of liquid cooled products is the ability to cool IT loads with elevated water temperatures. The relatively high inlet fluid temperatures for both Immersion cooling, and HP’s Rack Fluid Cooling has huge implications in reducing both energy and water use.

Shumate Engineering has developed a unique central plant design that eliminates the pitfalls of direct adiabatic cooling for high density mission critical applications. Using DLC technologies, this central plant solution can handle densities as high as 40 KW/cabinet and has a significantly lower PUE than any all-air system. Our solution uses existing, proven technology for all the components to serve Immersion Cooling, Rack Fluid Cooling or any similar type of DLC technology. We call this design Hybrid Dry Adiabatic Cooling (HDAC)[™].

Hybrid Dry Adiabatic Cooling (HDAC)™

Our solution for Hybrid Dry Adiabatic Cooling™ offers several advantages for data center operators that are evaluating how to integrate DLC equipment into their portfolio. Figure 1 illustrates a schematic diagram of the proposed HDAC central plant.



At the core of the system is the use of Adiabatic Fluid Coolers (AFC). An AFC is a heat rejection device that uses fans to induce air flow through a dry adiabatic pad and a closed loop coil to cool the fluid (water in this example). When outdoor ambient temperatures are below 81.5° F, all the heat transfer is dissipated with a dry coil. When outdoor ambient temperatures exceed 81.5° F, the adiabatic media is wet, which assists the dry coil to reject heat from the facility return water loop. The level of desired reliability and the load will determine the number of AFCs deployed. Utilizing elevated supply water temperature for the DLC equipment, the AFC can serve those units with zero water consumption for 94% of the year, based on typical weather data for the mid-Atlantic region. With the use of AFC equipment, there are a myriad of options to optimize energy consumption and reduce water use. The key is knowing the right balance to achieve a solution that optimizes energy and water use together. The supply water from the AFC is pumped to the facility with a number of parallel pumps. The required flow rate and desired reliability will determine the number of pumps deployed.

In a mission critical facility, we expect that an optimized Direct Liquid Cooled solution could account for as much as 80% of the total IT load. The remaining 20% of the load would still require some type of air-cooled solution. The advantage of the Hybrid Dry Adiabatic Cooling (HDAC) system is that we utilize the same elevated supply water temperatures from the AFC to pass through the evaporator bundle in the chilled water plant. The water-cooled chiller is used to create a secondary chilled water distribution loop at lower temperature, to serve any traditional air handling components in the data center. In this example, Computer Room Air-Handling Units (CRAHU's) would provide air circulation to the white space to supplement the DLC equipment. After the fluid leaves the air handling equipment, it then passes through

the condenser bundle in the chilled water plant and is combined with the return water from the fluid cooled equipment, before returning to the AFC's.

The chilled water supply to the CRAH equipment would be designed for 80° F to the equipment coils. The quantity of chillers selected is based on the desired level of reliability and the load. When operated at these entering water conditions, there is no chance of condensation at the coil since the fluid temperature will be above the dew point of the space. The system can be deployed to serve individual AHUs in a variety of configurations, including traditional downflow CRAHU's with a raised floor, horizontal CRAHUs hung above the hot aisle, or Fan Wall units. All of these air handling system types can be configured to utilize a hot aisle containment zone to maximize efficiency.

Current design standards for most hyperscale data centers include a 20° F delta-T between the supply air temperature and the hot aisle return air temperature. Anecdotal data indicates that the current densities of 11 kW/ cabinet have a higher temperature rise through the cabinet. Moving forward, designers and clients should consider increasing the delta-T in their design standards to reflect the actual operational temperatures. Even a 10% increase in air delta-T would result in a 10% reduction in fan motor kW. In addition, proper sizing of chilled water coils is another key factor to optimize energy performance. The old rule of using a standard coil face velocity of 500 feet per minute (FPM) is not recommended. Instead, the coil face velocity should be reduced to the lowest level that does not induce laminar flow.

Based upon our calculations, this central plant configuration offers tremendous energy savings and dramatically reduces water consumption as compared to direct adiabatic all-air cooled systems. We calculate the PUE for a hyperscale data center in the mid-Atlantic region from 1.028 to 1.062, depending on the ratio of air cooled versus fluid cooled equipment.

In conclusion we can state the following:

Direct Liquid Cooling provides one of the best opportunities for reducing PUE of new data centers and allows for higher rack density. Liquid cooling offers data center operators an opportunity to integrate a new generation of IT equipment. The two predominant DLC solutions, Immersion Cooling and Rack Fluid Cooling, are both poised to alter the landscape in the high-density mission critical environment. The industry adoption of cooling IT equipment with liquid at the source of heat is the perfect opportunity to implement Hybrid Dry Adiabatic Cooling.

Why HDAC matters:

Data centers are increasingly challenged by cost and availability of power, cost of land, and availability of cooling system make-up water and subsequent pressure from utilities, neighbors, and planning officials. An HDAC cooling approach that increases density – thus reducing required footprint, reduces overall power usage, and limits water results in a more environmentally responsible and financially competitive data center using less power per critical kW and significantly less water.

References:

1. Green Grid, Power Usage Effectiveness (PUE), "Harmonizing Global Metrics for Data Center Energy Efficiency", February 2, 2010.
2. HPE CRAY EX LIQUID-COOLED CABINET FOR LARGE-SCALE SYSTEMS: Copyright 2022 Hewlett Packard Enterprises Development LP
3. Penrod, J. (ed.), Shumate, D. (ed.), Weaver, J. (ed.): (2023), Direct Liquid Cooling and its Impact on the Design of Mission Critical Facilities. <https://www.shumateengineering.com>