The Future of Containment: Has Air Cooling Reached it’s Limits?

WHITEPAPER

SUBZERO ENGINEERING
By Gordon Johnson

INTRODUCTION

Today, many data centers are experiencing increasing power density per IT rack, rising to levels that just a few years ago seemed extreme and out of reach, but today are considered both common and typical while simultaneously deploying air cooling.

This increase has left data center designers and managers wondering if air cooled ITE, along with containment used to separate the cold supply air from the hot exhaust air, has finally reached its limits and if liquid cooling is the long-term solution. It's not a simple yes or no answer. Moving forward it's expected that data centers will be transitioning from 100% air cooling to a hybrid of air and liquid cooling, with all new and existing air-cooled data centers requiring containment. Additionally, even those moving to liquid cooling may still need containment depending on the technology deployed.

Why is the debate of air versus liquid cooling such a hot topic in the industry right now? To answer this question, we need to understand what's driving the need to consider liquid cooling in some data centers. Are there other options? How can we evaluate these options while continuing to cool with air? This paper will answer these and many more questions in the topics noted to the right.

This paper will discuss these topics:

- **Can Air and Liquid Cooling Coexist**
- **What Do Server Power Trends Reveal**
- **What's Driving Server Power Growth**
- **Additional Options for Air Cooling**
- **Closed Loop Versus Open Loop Cooling Systems**
- **Liquid Cooling for Air-Cooled ITE**
- **Air-Assisted Liquid Cooling (AALC)**
- **Liquid Cooled ITE**
- **Immersion Liquid Cooling**

CAN AIR AND LIQUID COOLING COEXIST?

We've been in this position before, with air and liquid cooling successfully coexisting while removing substantial amounts of heat via intra-board air to water heat exchangers. This process continued till the industry shifted primarily to CMOS technology in the 1990's, and we've been using air cooling in our data centers ever since.

With air cooling being the primary source used to cool data centers, ASHRAE (American Society of Heating, Refrigeration, and Air Conditioning Engineers) has worked towards making this technology as efficient and sustainable as possible. Since 2004 they've published a common set of criteria for cooling IT servers with the participation of ITE and cooling system manufacturers entitled "TC9.9 Thermal Guidelines for Data Processing Environments".

ASHRAE has focused on efficiency and reliably cooling the ITE in the data center. Several revisions have been published with the latest being released in 2021 (revision 5). This latest generation TC9.9 highlights a new class of high-density air-cooled ITE (H1 class) which focuses more on cooling high-density servers and racks with a trade off in terms of energy efficiency due to lower cooling supply air temperatures recommended to cool the ITE (**see FIGURE 1**).

As to the question on whether or not air and liquid cooling can coexist in the data center white space, it's already been done for decades, and moving forward experts in both cooling technologies expect to see these two cooling technologies coexisting for many years to come.[1]

Classes	°C	°F
Recommended (Suitable for Classes A1 to A4)	18 to 27° C	64.4 to 80.6° F
A1	15 to 32° C	59 to 89.6° F
A2	10 to 35° C	50 to 95° F
A3	5 to 40° C	41 to 104° F
A4	5 to 45° C	41 to 113° F
H1 Recommended	18 to 22° C	64.4 to 71.6° F
H1 Allowable	15 to 25° C	59 to 77° F

FIGURE 1 Source: ASHRAE TC9.9 - Thermal Guidelines 5th Edition 2021

WHAT DO SERVER POWER TRENDS REVEAL

It's easy to assume that when it comes to cooling that a one size data center will fit all designs in terms of power consumption both now and in the future, but that's not accurate. It's more important to focus on the actual workload for the data center that's we're designing or managing.

For example, the chart below (**see FIGURE 2**) on the vertical axis shows heat load (kW per rack) versus year of product introduction. In the past it was a common assumption with air cooling that once you went above **25 kW** per rack it was time to transition to liquid cooling. But the industry has made some changes with regard to this number as will be described later in this paper that is enabling data centers to cool up to and even exceed **35 kW** per rack with traditional air cooling.

In 2022 the average rack density was approximately **10.5 kW**, which is similar to what **FIGURE 2** shows while highlighting that all data centers are different in terms of workloads, growth rates, and power consumption. Scientific data centers, which include largely GPU driven applications like machine learning AI and high analytics like crypto mining, are the areas of the industry that typically are transitioning or moving towards liquid cooling.

But if you look at some other workloads like the cloud and most businesses, the growth rate is rising but it still makes sense for air cooling in terms of cost. The key is to look at this issue from a business perspective, what are we trying to accomplish with each individual data center.[1]

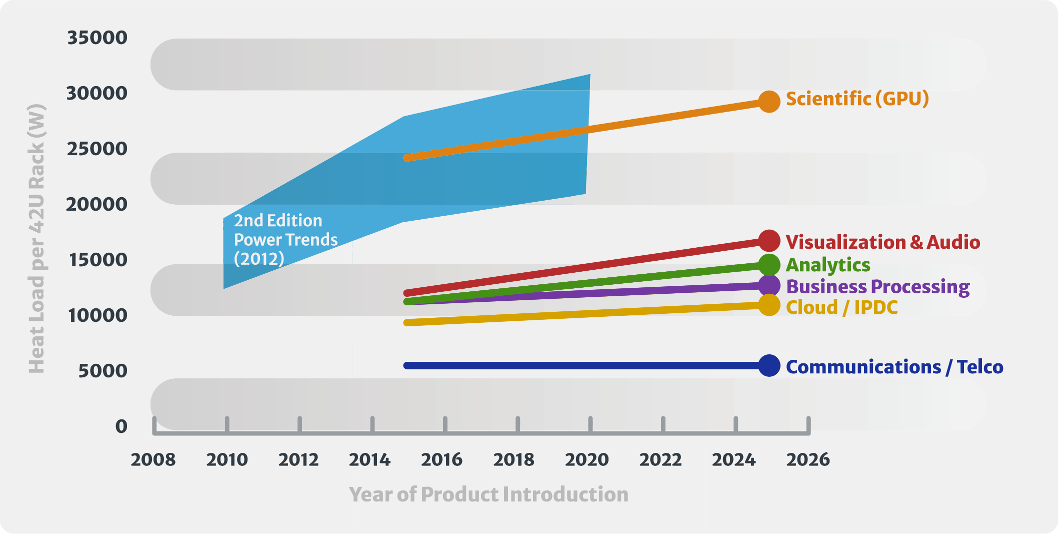


FIGURE 2 Source: ASHRAE 2018 - IT Equipment Power Trends, 3rd Edition Datacom Book2

WHATS DRIVING SERVER POWER GROWTH

The chart below (**see FIGURE 3**) displays or represents commercial CPUs on the market each year starting in 2000. Up to about 2010 we just had single-core processors, but then transitioned to multi-core processors which meant more cores on the processor, but as shown in the chart there still was a relative flat power consumption for CPUs. That enabled server manufacturers to concentrate on lower airflow rates for cooling ITE which resulted in better overall efficiency.

But **FIGURE 3** also highlights that around 2018, multi-core processors became the norm and now with these reaching their performance limits, the only way to continue to achieve increased performance is by increasing power consumption. Server manufacturers have been packing in as much as they can to servers, but because of CPU power consumption in some cases we're having difficulty removing the heat with air cooling. This is one reason behind the need for liquid cooling for some business cases.

Another reason is server manufacturers have been increasing the temperature delta across servers for several years now, which again has been great for efficiency since the higher the temperature delta the less airflow (CFM) that's needed to remove the heat from the ITE. However, server manufacturers are reaching their limit with increasing the temperature delta, resulting in having to increase the airflow, especially for high-density servers, to keep up with increasing power consumption. This is another reason behind the need for liquid cooling for some business cases.[1]

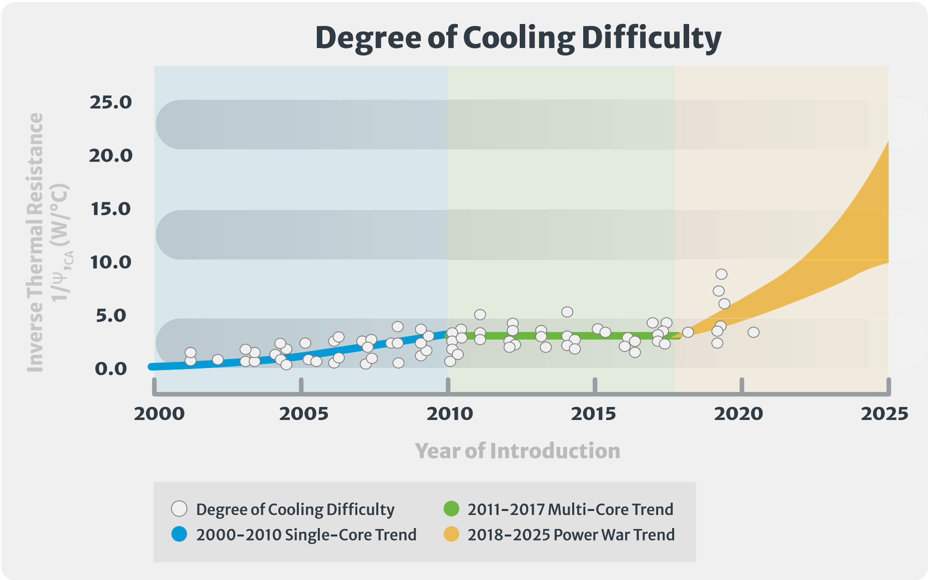


FIGURE 3 Source: ASHRAE 2021 - Emergence and Expansion of Liquid Cooling in Mainstream Data Centers

ADDITIONAL OPTIONS FOR AIR COOLING

There are several options that many in the industry are successfully using to cool up to and even exceed **35 kW per rack** with traditional air cooling. These options start with deploying either full cold or hot aisle containment. If no containment is used typically rack density should be no higher than **5 kW per rack** with additional supply airflow needed to compensate for recirculation air and ITE hot spots. Other options include larger cold aisles (6' instead of 4'), opening up the plenum ceiling above the HAC as much as possible (100% open is optimal), and increasing the width of the hot aisles when applicable.

What about lowering temperatures? As mentioned earlier, the industry is now doing just that by lowering temperatures for air cooling in data centers. In 2021, ASHRAE released their 5th generation TC9.9 which highlighted a new class of High-Density Air-Cooled IT equipment called the H1 class, and this new class of ITE will need to be cooled to more restrictive supply temperatures than the previous class of servers (**see FIGURE 1**).

At some point, high-density servers and racks will need to transition from air to liquid cooling, especially with CPUs and GPUs expected to exceed 500 watts per processor or higher in the next few years. But this transition is not automatic and isn't going to be for everyone, both now and in the future.

Liquid cooling is not going to be the ideal solution or remedy for all future cooling. Instead, the selection of liquid cooling instead of air cooling has to do with a variety of factors, including specific location, climate (temperature/humidity), power densities, workloads, efficiency, performance, heat reuse, and physical space.[2]

Again, this really highlights the need for data center designers and managers to take a holistic approach to cooling their data centers. It will not and should not be an approach where we're considering only air or only liquid cooling moving forward. Instead, the key is to understand the trade-offs of each cooling technology and what makes the most sense for each individual business case.

CLOSED LOOP VERSUS OPEN LOOP COOLING SYSTEMS

Regardless of which technology is being used, cooling systems fall into one of two categories – Closed Loop Systems versus Open Loop Systems (**see FIGURE 4**).

First, with an Open Loop System it's important to understand that either cold or hot aisle containment is almost always needed since the airflow to and from the racks and cooling units is affected by and affects the temperature and humidity in the room.

Second, with a Closed Loop System, containment is almost never needed since the cooling loop acts as its own micro climate within a single IT rack or multiple racks in a row. This means that no air, cold or hot, is rejected into the surrounding space as it only cools the environment within the rack and isn't affected by ambient air (nor does it affect that air).

Liquid cooling systems are either closed or open loop systems and generally fall into four categories that the following sections will discuss and also explain whether containment is needed or not.

The four categories of Liquid Cooling

1 **Liquid Cooling for Air-Cooled ITE**

2 **Air-Assisted Liquid Cooling (AALC)**

3 **Liquid Cooled ITE**

4 **Immersion Liquid Cooling**

OPEN LOOP SYSTEM

Open Loop Systems **almost always need containment** since the airflow to and from the racks and cooling units is affected by and affects the temperature and humidity in the room.

CLOSED LOOP SYSTEM

Closed Loop Systems **almost never need containment** since the cooling loop acts as its own microclimate within a single IT rack or multiple racks in a row. Typically, no cold or hot air is rejected into the surrounding space of the data center.

FIGURE 4 Closed Loop versus Open Loop System

LIQUID COOLING FOR AIR-COOLED ITE

The ITE for this technology is standard unmodified equipment. The process starts at the processor chip and the heat is typically transferred to an attached air-cooled heat sink. Next, the heat is transferred back to the cooling units or to CCC (Close Coupled Cooling).[3] With the latter, since air usually cools better the closer it is to the ITE, the result is air can be used to cool higher density racks without having to deploy the new H1 Class ITE with its lower supply temperature requirements, so efficiency isn't sacrificed. CCC almost always needs some form of containment since the airflow to and from the racks and cooling units is affected by and affects the temperature and humidity in the room.

IRC (In-row cooling) is an example of CCC where the cooling, climate control, and heat removal is brought directly to the data center rows. IRC can be the main form of cooling for high-density racks, or it can be used to supplement existing cooling for higher density installations within the data center. Either CAC or HAC will be needed when using IRC.

Another form of CCC are RDHx (Rear Door Heat Exchangers) which can remove 100% of the heat and can efficiently cool high-density racks. The two types are passive RDHx which utilize existing server fans to push the heated exhaust air across a liquid-filled coil that absorbs the heat before the air is returned to the data center. The second type are active RDHx which have additional fans that pull the exhaust air across the liquid-filled coil which again removes the heat before the air is returned to the data center.

FIGURE 5a
IRC (In-Row Cooling)
- Bing Images

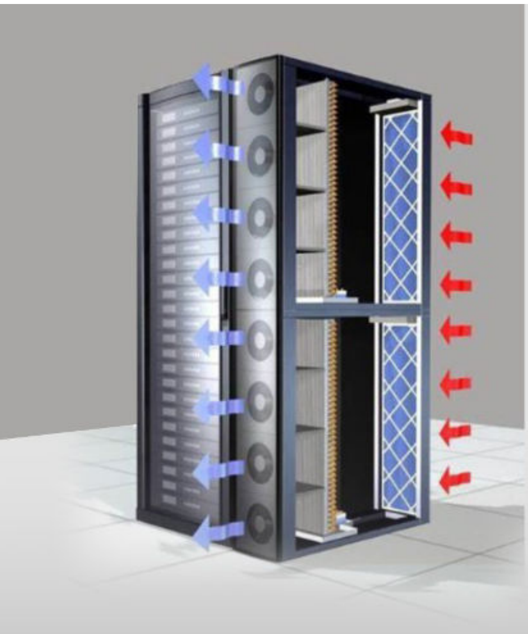
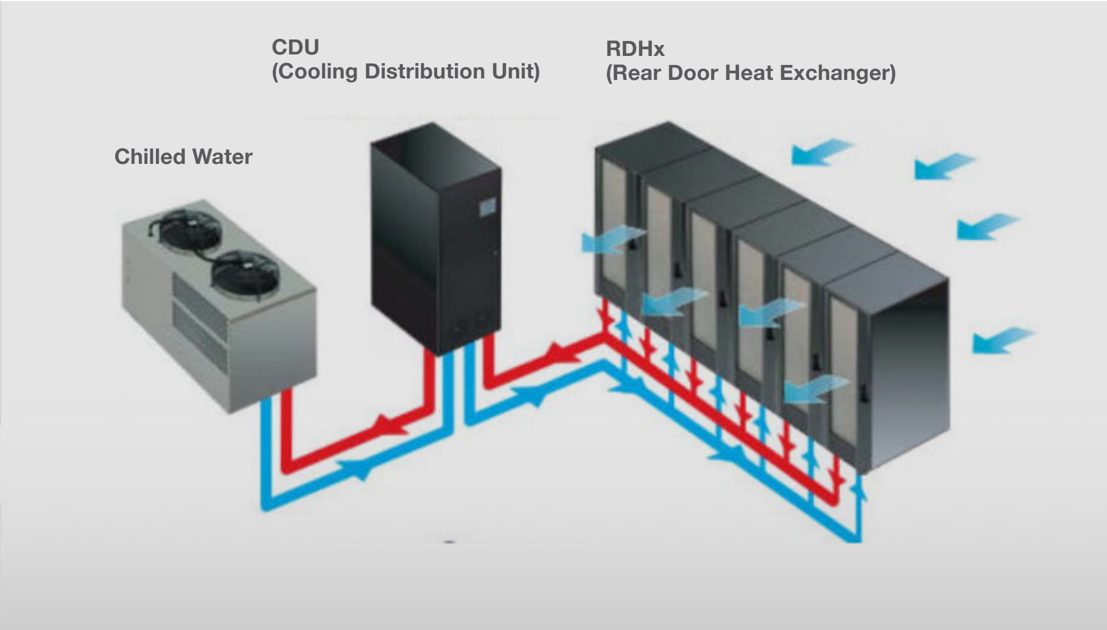


FIGURE 5b
Rear Door Heat Exchanger (RDHx)
- Bing Images



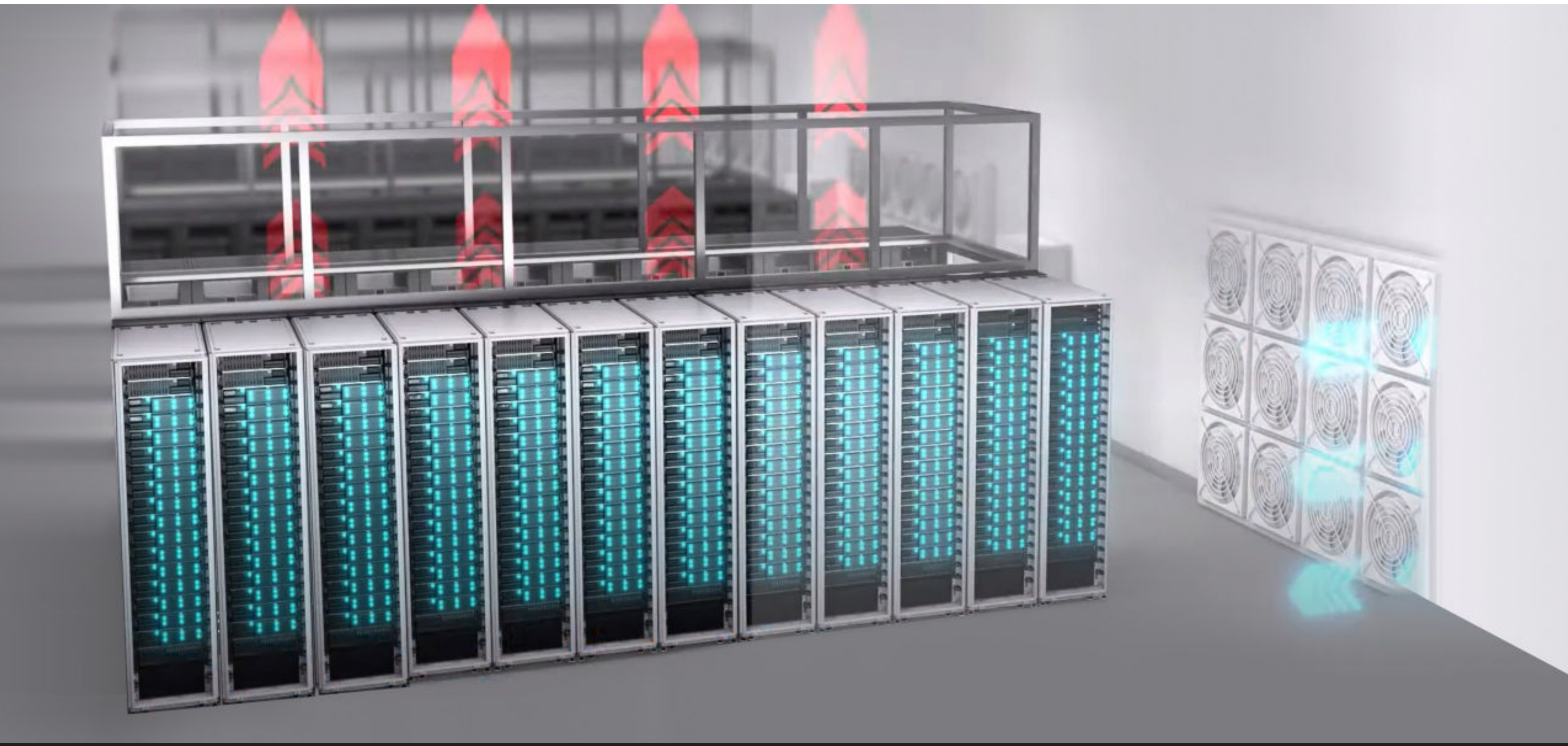
If the data center has all RDHx, no containment is needed. If the data center has existing air-cooled ITE and those racks currently have CAC, then some additional containment will be needed to prevent the exhaust air from those racks from entering the racks with RDHx. If the existing air-cooled racks already have HAC, then no additional containment is needed.

AIR ASSISTED LIQUID COOLING (AALC)

This technology can assist data center designers and managers that need to cool high-density racks but are not ready or don't want to make the shift to liquid cooling for various reasons. AALC is a good upgrade for extending the life of existing air-cooled and new HPC facilities since it serves as a transitional shift to liquid cooling while still using traditional air cooling (CRACs, CRAHs, Fan Walls, etc.) to flood the data center with cold supply air.

Several large hyperscale customers are already moving towards AALC since it can cool between **45 kW and up to 55 kW per rack** in some cases. An internal RPU (Reservoir Pump Unit) and heat exchanger assist the existing air cooling to cool these higher densities, without the need to run piping with liquid to and from the racks. HAC will be needed when using AALC.

FIGURE 6
Air-Assisted Liquid Cooling (AALC) Courtesy Photo from Delta



LIQUID COOLED ITE

Direct liquid cooling, also known as DTC (Direct to Chip) utilizes pipes that deliver liquid directly onto a cold plate which sits atop the CPU/GPU to draw off the heat. The extracted heat is transferred to a chilled-water loop to be expelled into the outside atmosphere. Typically, a portion (50-75%) of the heat is transferred to a liquid loop while the remainder of the heat is removed via traditional room-based cooling units (CRACs, CRAHs, Fan Walls, etc.).

Since traditional air cooling is needed to remove the remaining heat generated by power supplies, memory, other circuit boards, etc., either CAC or HAC will be needed for these deployments.

FIGURE 7
DTC (Direct to Chip) Photo
Direct to Chip Liquid Cooling - Bing images

IMMERSION LIQUID COOLING

Immersion Liquid Cooling submerges the ITE in a special fluid or dielectric liquid, which engulfs the chassis and all the components, thus transferring virtually all the heat to the fluid or liquid. Like other liquid cooling methods, the underlining principle behind immersion cooling is that liquid is a much better heat conductor than air. With immersion-based systems the liquid is in direct proximity to the heat (or as close as you can get). This is an example of a true Closed Loop System, so no containment will be needed.

Immersion cooling can be single-phase or two-phase. With single-phase immersion cooling, the coolant is continuously circulated and cooled to dissipate the heat. In a two-phase immersion cooling system, a coolant with a low boiling point is used. When the coolant boils, it turns to vapor and rises to the top of the tank or chassis lid, where it's cooled and condensed back to liquid.

Immersion cooling is very efficient but can also be complex to operate and maintain due to fluid types and operating temperatures.

CONCLUSION

Once again, how do we move forward? There's no clear-cut answer, the choice will really be dependent on each data center facility. Experts in both air and liquid camps tend to agree that future data centers will need both high-density and regular density equipment, it will not be a one size fits all solution.

Data center designers and managers will also need to design data centers while separating both high-density and low-density equipment, this is how we'll drive future efficiency, not just one type of cooling versus the other. In addition, understanding the TCO (Total Cost of Ownership) of cooling the data center is essential. The decision to deploy either air cooling or liquid cooling should always be determined by TCO. This is what we'll see more in the future, each facility will need to examine what they're doing now and want to do with their data center in the future.[1]

In summary, ITE is changing and evolving and will continue to do so. ITE power and cooling has and will continue to push the limits due to packaging density and HPC. Additionally, air cooling and liquid cooling experts agree that both technologies will coexist and share the white space for many years to come, with future data centers transitioning from 100% air cooling to a hybrid of air and liquid cooling driven by energy consumption concerns and sustainability goals.

And finally, with air and liquid cooling coexisting moving forward, containment will continue to be the fundamental starting point for maximizing available cooling capacity and energy efficiency. A well designed and implemented containment system is and will continue to be an essential element to support current and future high-density air cooled ITE.[3]

ABOUT THE AUTHOR

Gordon Johnson is the Senior CFD Engineer at Subzero Engineering and is responsible for planning and managing all CFD related jobs in the U.S. and worldwide. He has over 30 years of experience in the data center industry which includes data center energy efficiency assessments, data center design, CFD modeling, and disaster recovery. He is a certified U.S. Department of Energy Data Center Energy Practitioner (DCEP), a certified Data Centre Design Professional (CDCDP), and holds a Bachelor of Science in Electrical Engineering from New Jersey Institute of Technology. Gordon also brings his knowledge and ability to teach the fundamentals of data center energy efficiency to numerous public speaking events annually, white papers, and industry leading researched articles worldwide.

REFERENCES

1. **DCD Keeping IT Cool Broadcast – Understanding the Current State of Affairs for Air Cooling: Has it Had its Day?**
   Posted on August 3, 2022 by Dustin Demetriou
   www.datacenterdynamics.com/en/broadcasts/kitc-air/2022/episode-two/

2. **Data Centers Predicted to Drive Liquid Cooling Market**
   Posted on December 12, 2022 by Sharon Fisher
   www.datacenterknowledge.com/power-and-cooling/data-centers-predicted-drive-liquid-cooling-market

3. **Data Center Frontier Special Report – High Density IT Cooling Breaking the Thermal Boundaries**
   Posted on June 24, 2022 by Julius Neudorfer
   www.datacenterfrontier.com/special-reports/article/11427279/pushing-the-boundaries-of-air-cooling-in-high-density-environments