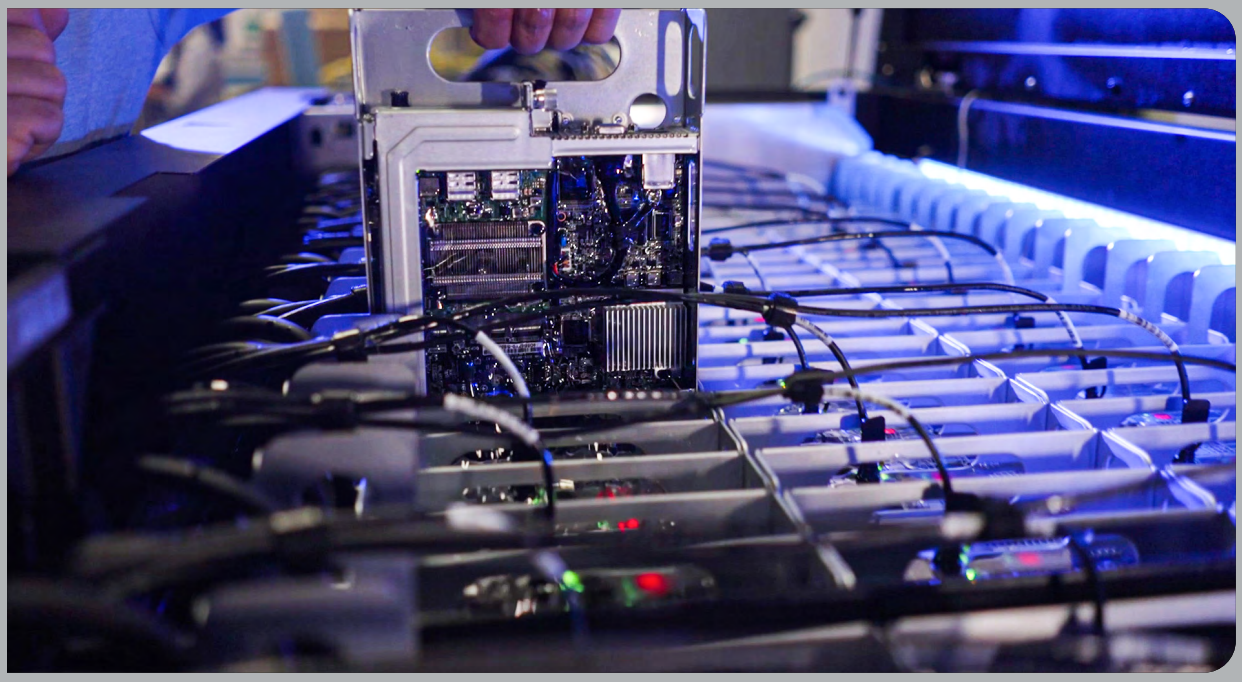
DATA CENTER Frontier

SPECIAL REPORT

Liquid Immersion Cooling Will Save The Data Center: Critical Design and Infrastructure Updates

Liquid Cooling Paper Update: Considerations in the AI Era



Author: Bill Kleyman, Contributing Editor, Data Center Frontier

brought to you by

Hypertec

Contents

Liquid Immersion Cooling Will Save The Data Center: Critical Design and Infrastructure Updates

SUMMARY

Summary: Transforming every data center into an AI data center is not a matter of if but when. The critical updates in the field of liquid cooling since the last paper have accelerated this transformation. New innovations around manageability, ease of service, density, and new applications have propelled liquid cooling into the forefront of data center design. This is not a trend but a necessity.

The exponential growth in computational demands of large-scale (and small-scale) AI systems requires an immediate paradigm shift in data center design.

To that point, we have all become users AI. None of this is a fad. AI is poised to drive a 160% increase in data center power demand and will use 8% of US power by 2030, compared with 3% in 2022. In terms of infrastructure design to support this massive growth, one fact is clear: Our data centers have run out of air. Leaders must turn to some liquid cooling solution to dissipate all of this newly generated heat and create greater density levels. There are also massive cost considerations, and it's a big reason why many leverage liquid immersion cooling to reduce cooling OPEX by over 90% and reduce CAPEX building costs by more than 50%.

This paper will explore critical updates to our previous special report. Specifically, we will cover the existential situation facing many data center leaders. It's official; we have entered the *Era of Liquid Cooling*. The special report will explore what this means; we'll dispel some new liquid immersion cooling myths and discuss how you can adapt these technologies to your facilities.

Introduction

The second ChatGPT hit the mainstream, a seismic shift occurred in how humanity interacts with data. Even though many in the data center industry didn't realize it, their world was about to shift as well. To look ahead, we'll need to look backward just a bit.

Eight years ago, in the first AFCOM State of the Data Center report, we asked the all-important question about rack density. Respondents indicated that their average density was 6.1 kilowatts per rack in that report. Given the typical types of workloads data centers support, this metric is pretty expected. While high-density applications were a thing, most data centers were still running very traditional applications like e-mail servers, databases, and other business-critical services.

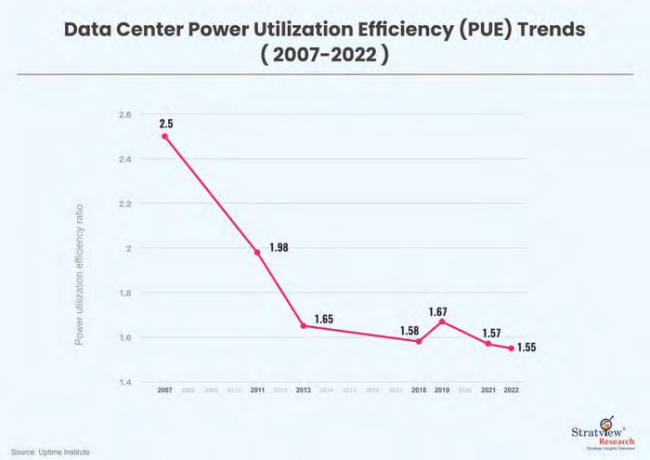
Over the past year, everything has changed. In the 2024 report, respondents indicated that the rack's average density had increased to 12 kilowatts. Most respondents (60%) are actively working to increase density in their racks (58%), primarily by improving airflow, followed by containment (42%) and liquid cooling (40%). What was driving all of this? AI.

Per the report, most respondents (53%) believe new AI workloads (generative AI) will "definitely" increase capacity requirements for the colocation industry.

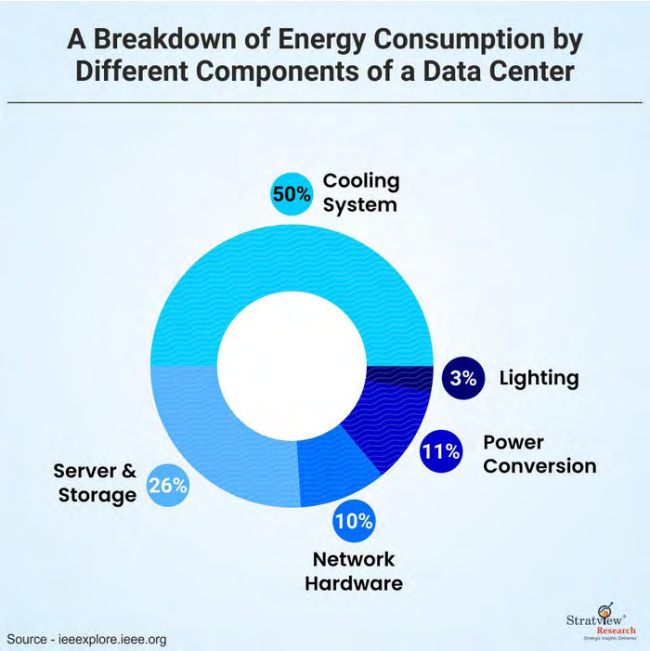
Here's the crazy part: Even though density has doubled, at least per the report, it's still insufficient to support AI and high-density architecture. In perspective, a single node of NVIDIA's DGX H100 cards (eight cards per node) will consume between 10.2 kilowatts per rack. Based on our findings, a traditional data center could only support one of these high-end units in their rack despite growth in density. This rapid pace of evolution has become the driving force of innovation in our industry. The most significant difference is that it's happening incredibly fast. While there isn't current support for immersion cooling of H100s, there are already designs supporting immersion cooling GPUs. We'll discuss this in a later section.

This is where liquid cooling must come into play. Because of AI, the amount of power our systems require will only continue to grow. In 2023, data centers across the globe consumed 7.4 GW of power, a 55% increase from the 4.9 Gigawatts in 2022, according to Cushman & Wakefield. That number is already forecast to grow significantly through 2030.

Massive strides have been made in terms of efficiency. The PUE trends have indicated that we are indeed using our data center resources more wisely.



However, the amount of power consumed by servers and cooling is still quite a bit.



As we examine emerging use cases (like AI and HPC), we must understand their impacts on cooling systems, networks, storage, and compute. This involves everything from liquid cooling GPUs to deploying a complete, all-in-one compute, storage, and network architecture with liquid cooling built-in.

This next point is critical: *Liquid immersion cooling has become a staple design mechanism to support HPC and new AI use cases.*

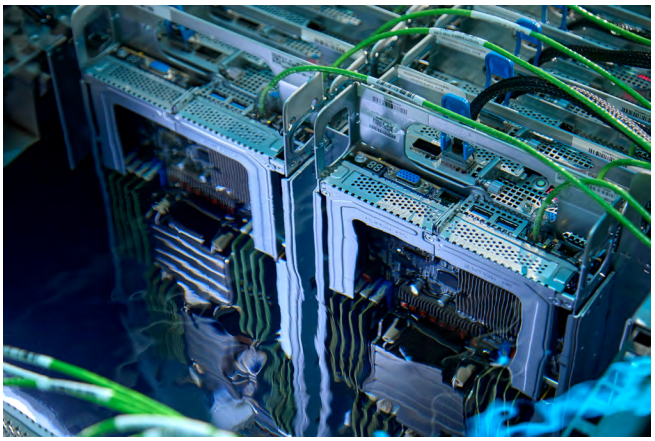
This means new imperatives exist in how we cool, maintain, and optimize our data center facilities.

LIQUID IN THE DATA CENTER: YES, IT'S A GOOD THING

Liquid immersion cooling isn't new. However, recent advancements have allowed more data centers to adopt it. Why is this happening? AI and HPC systems require more resources and consume more power. The growing emphasis on reducing data center power consumption is driving the market's growth.

Increasing concerns about rising power consumption in data centers that use air-cooled technologies have significantly increased the need for liquid immersion cooling and immersion solutions. Data center companies actively invest in new liquid immersion cooling solutions to reduce operating expenses (OPEX). However, they are still struggling to develop data center infrastructure that requires minimal maintenance and less power. Again, working with liquid immersion cooling solutions that are easy to service and have simple maintenance cycles is definitely the way to go.

Remember, immersion cooling technology can reduce upfront data center costs by 60%, cooling power consumption by 95%, server power consumption by 10–20% Also, more racks at data center sites will



increase rack density, thus making companies more inclined to use data center liquid immersion solutions over air.

Finally, emerging **advanced technologies** are a primary trend in the market. Advanced technology like machine learning, artificial intelligence, and blockchains are increasingly being adopted and used. Powerful new hardware for AI workloads can also produce high flux, wherein the power usage in a rack surges rapidly as the hardware begins a new workload, which can be challenging to manage using the traditional air method.

Let's pause here and focus on you, the data center leader. Demand for new data centers will continue to grow. That part is prominent. However, the use cases tenants and clients ask for will be fundamentally different. Can today's air-cooled ecosystems keep up? Immersion cooling has pushed boundaries on efficiency and density. Today, these solutions are a bridge to deliver the most advanced workloads to your customers.

With that, let's examine some of the biggest updates in the liquid immersion cooling industry and how these new innovations will impact your data center.

Section 1 – Liquid Cooling: A year in review

Between 2005 and 2010, liquid cooling entered the data center with chilled doors. From there, looking into 2010 and beyond, liquid cooling was used in high-performance computing (HPC) environments and designs featuring direct contact and total immersion liquid immersion cooling solutions. But we're not here to look at the past decade. We're here to look at what happened alone in the last year. Buckle up.

LIFE BEFORE AND AFTER AI/CHATGPT

Introducing AI technologies, particularly generative AI models like ChatGPT, has marked a seismic shift in the data center industry. Before AI became mainstream, data centers were primarily designed to support traditional applications such as email servers, databases, and basic web hosting services. These applications typically required low to moderate levels of computational power, and the corresponding cooling requirements were relatively modest. Air cooling, with average rack densities of around 6.1 kilowatts, was sufficient to maintain the operational integrity of these data centers.

However, the rapid adoption of AI and machine learning applications has dramatically changed this landscape. According to a recent AFCOM State of the Data Center report, the average rack density has doubled to 12 kilowatts, driven primarily by the increased computational demands of AI workloads. The traditional air-cooled infrastructure cannot handle these high-density environments, leading to overheating issues, increased energy consumption, and significant operational costs.

In this new era, liquid cooling has become an option and a necessity. The shift from air to liquid cooling represents a fundamental change in how data centers are designed and operated. Liquid immersion cooling systems, which involve submerging hardware components in non-conductive liquids or directly applying coolant to heat-generating components, offer superior cooling efficiency and energy savings. As AI continues to drive up the computational and cooling demands of data centers, liquid immersion cooling will play an increasingly vital role in maintaining the performance and reliability of these critical infrastructures.

HOW LIQUID IMMERSION COOLING WILL SAVE THE DATA CENTER

The limitations of traditional air-cooling systems have become more apparent as data centers strive to meet the demands of modern workloads. AI and HPC applications generate significantly more heat than traditional workloads, and air cooling alone is no longer sufficient to dissipate this heat effectively. This has led to increased energy consumption, higher operational costs, and the need for more sophisticated cooling solutions.

Liquid cooling, particularly immersion cooling, addresses these challenges by offering a more efficient heat dissipation method. Immersion cooling involves submerging servers in a dielectric fluid, which absorbs the heat generated by the components and transfers it away from the hardware. This method reduces the temperature of the components more effectively than air cooling and allows for higher densities within the data center. *Liquid immersion cooling significantly reduces the facility's overall footprint and operational costs by enabling data centers to operate at higher densities without overheating.*

Moreover, liquid immersion cooling systems can reduce cooling operational expenditures (OPEX) by up to 95% compared to traditional air-cooling systems. This reduction in OPEX, combined with the potential for increased rack density, makes liquid immersion cooling an essential technology for the future of data centers. As data centers continue to scale up to meet the demands of AI and HPC applications, adopting liquid immersion cooling will be critical to maintaining efficiency, reliability, and sustainability.

SIGNIFICANT TRENDS IN DATA CENTER COOLING, POWER, AND AI

The shift towards liquid immersion cooling is part of a broader trend in the data center industry, driven by the increasing demands of AI and HPC applications. Several key trends are shaping the future of data center design and operation:

- **Cooling:** One of the most significant trends in the data center industry is the need for more efficient cooling solutions. As computational power increases, so does the heat generated by data center hardware. Liquid immersion cooling offers a more effective solution than air cooling, enabling data centers to maintain optimal temperatures even as they scale up to meet the demands of AI and HPC workloads.

- **Power Consumption:** The rise of AI and HPC has also led to a significant increase in power consumption within data centers. According to industry reports, AI workloads are expected to drive a 160% increase in data center power demand by 2030, with AI alone consuming 8% of US power. Liquid immersion cooling helps mitigate this increase in power consumption by improving the efficiency of the cooling process and reducing the overall energy required to maintain data center operations.

- **AI and HPC:** The growing importance of AI and HPC drives the adoption of new technologies and solutions within data centers. These applications require more computational power and sophisticated cooling solutions to manage the heat generated by high-density hardware. Liquid immersion cooling is becoming increasingly essential for data centers that support AI and HPC workloads, as it can handle high-density environments and reduce energy consumption.

NEW TYPES OF BUSINESS DEMANDS AND EFFICIENCY REQUIREMENTS

The shift towards AI and HPC drives new business demands within the data center industry. Enterprise customers increasingly seek data center partners who can provide the infrastructure and services needed to support these advanced applications, leading to greater efficiency, scalability, and sustainability requirements.

One of the most significant trends in this area is the growing demand for repatriation—bringing workloads back from public clouds to on-premises or colocation data centers. In the latest AFCOM State of the Data Center Report, most respondents (71%) have seen a migration of workloads from the cloud back to on-premise data centers or colocations. For those data center operators that saw repatriation, 71% have seen an increase in the power load demand, including 22% who report a significant impact. As businesses seek greater control over their data and IT infrastructure, they turn to data centers offering the high-density, efficient cooling solutions needed to support AI and HPC workloads.

In response to these demands, data centers are placing a greater emphasis on efficiency, both in terms of energy consumption and space utilization. Liquid immersion cooling systems enable data centers to meet these efficiency requirements by reducing energy consumption, maximizing rack density, and minimizing the physical footprint of the facility. As a result, component cooling and immersion cooling are becoming essential components of modern data center design, enabling facilities to meet the growing demands of their customers while maintaining operational efficiency and sustainability.

Section 2 – Immersion Liquid Immersion Cooling: The HPC and AI Imperative

FUNDAMENTAL CHANGES IN THE DATA CENTER BUSINESS

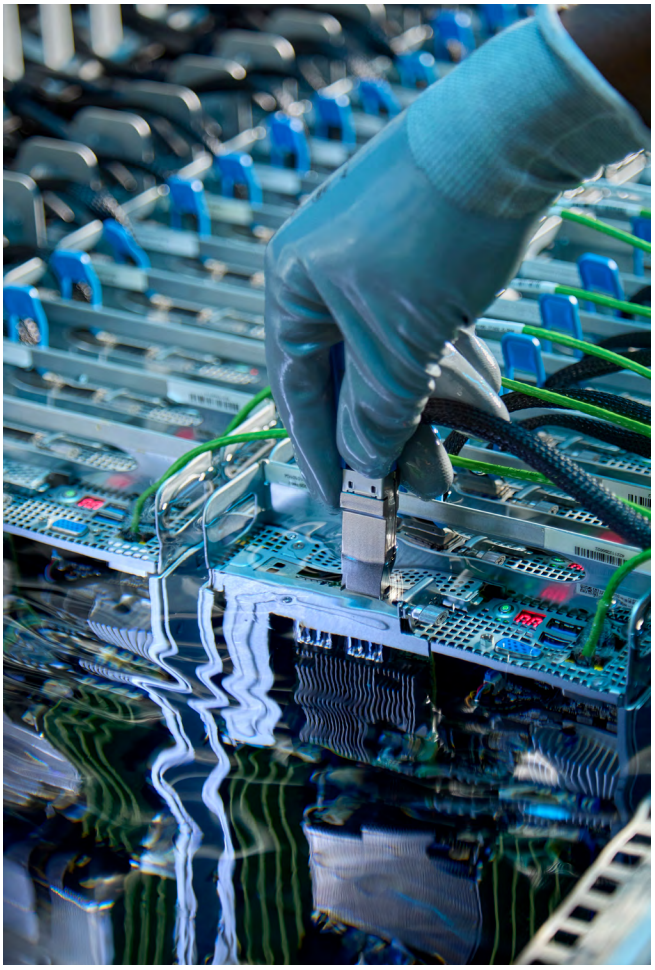
The data center industry is undergoing a significant transformation, driven by the increasing demands of AI and HPC applications. These applications require more computational power, generate more heat, and consume more energy than traditional workloads, necessitating a fundamental shift in how data centers are designed and operated.

One of the industry's most significant changes is the growing adoption of liquid immersion cooling. Immersion cooling, which involves submerging servers in a dielectric fluid, offers several advantages over traditional air cooling. It provides more efficient heat dissipation, increases densities, and reduces energy consumption, making it an ideal solution for data centers supporting AI and HPC workloads.

Several factors are driving the adoption of immersion cooling. First, the temperature fluctuations caused by AI workloads can be challenging to manage with traditional cooling methods. Immersion cooling provides a more stable and consistent cooling environment, essential for maintaining AI applications' performance and reliability. Second, HPC applications, which require large computational power, generate significant heat that cannot be effectively dissipated with air cooling alone. Immersion cooling provides a more efficient solution, enabling data centers to support these high-performance workloads without the risk of overheating.

DATA CENTER DESIGN CONSIDERATIONS

The shift towards immersion cooling also requires reevaluating data center design considerations. Traditional air cooling methods, which rely on fans and air circulation to dissipate heat, are no longer sufficient to meet the demands of modern data centers. New design considerations must be considered as data centers scale up to support AI and HPC applications.



One of the most critical considerations is the power implications of liquid immersion cooling. While immersion cooling systems may have a higher initial capital expenditure (CAPEX), they offer significant operational savings over time. Liquid immersion cooling systems reduce energy consumption, lower cooling costs, and enable higher densities, significantly reducing the total cost of ownership (TCO). As a result, data centers must carefully evaluate the long-term benefits of liquid immersion cooling when making design decisions.

Another important consideration is the cooling solution's scalability. The cooling system must scale with the facility as data centers grow and evolve.

Immersion cooling offers a scalable solution that can be adapted to meet the changing needs of the data center, making it an ideal choice for facilities that anticipate future growth.

CHALLENGES IN AIR COOLING SPECIFIC WORKLOADS

As the demands on data centers continue to increase, the limitations of air cooling are becoming more apparent. Let's be clear: there is nothing wrong with air cooling as long as it is used properly. The challenge is that with greater levels of density of AI and HPC, air cooling is less efficient at managing the heat generated by those workloads, leading to several challenges:

- **Scale:** As data centers scale up to support larger workloads, the limitations of air cooling become more pronounced. Immersion cooling offers a more scalable solution to handle the increased heat generated by high-density deployments.

- **Capacity:** Air cooling faces a significant challenge in efficiently cooling larger capacities, which immersion cooling effectively addresses. By providing more efficient heat dissipation, immersion cooling enables data centers to support higher capacities without the risk of overheating.

- **Efficiency and Performance:** Air cooling is less efficient at maintaining optimal operating temperatures, which can negatively impact the performance and reliability of data center hardware. On the other hand, immersion cooling provides more consistent and efficient cooling, improving overall system performance and reducing the risk of hardware failure.

While there may be a few challenges to work through, there are pre-existing liquid immersion cooling myths that have been well covered and reviewed. Understanding how far immersion cooling has come and its capabilities is important.

Mythbusting Immersion Cooling

Despite the growing adoption of immersion cooling, several myths and misconceptions persist about the technology. It is important to address these myths to provide a clear understanding of the benefits and challenges associated with immersion cooling:

- **Cable Wicking:** A common concern is that liquids could wick up cables and damage the hardware. However, proper design and cable management in immersion systems prevent this issue.

- **High-Pressure Systems:** Some believe immersion cooling systems require high-pressure environments, making them more complex and difficult to manage. In reality, immersion systems do not require high pressure and are designed to operate safely and efficiently at standard pressure levels.

- **Limited Use in Traditional Data Centers:** While immersion cooling has traditionally been used in specialized environments, such as HPC and AI data centers, it is also increasingly being adopted in more traditional data centers. As the benefits of immersion cooling become more widely recognized, its use is expected to grow across a broader range of data center environments.

- **Water Treatment Requirements:** Unlike water cooling systems, which require regular water treatment to prevent corrosion and scale formation, immersion cooling systems use dielectric fluids that do not promote these issues. This reduces maintenance requirements and simplifies the cooling system's operation.

- **Combustible Oils:** Another myth is that the oils used in immersion cooling are combustible, posing a safety risk. Modern dielectric fluids used in immersion cooling are non-combustible, making them safe for data centers.

- **Material Compatibility:** Concerns about material compatibility are also common but advances in materials science have addressed these issues. Most components used in

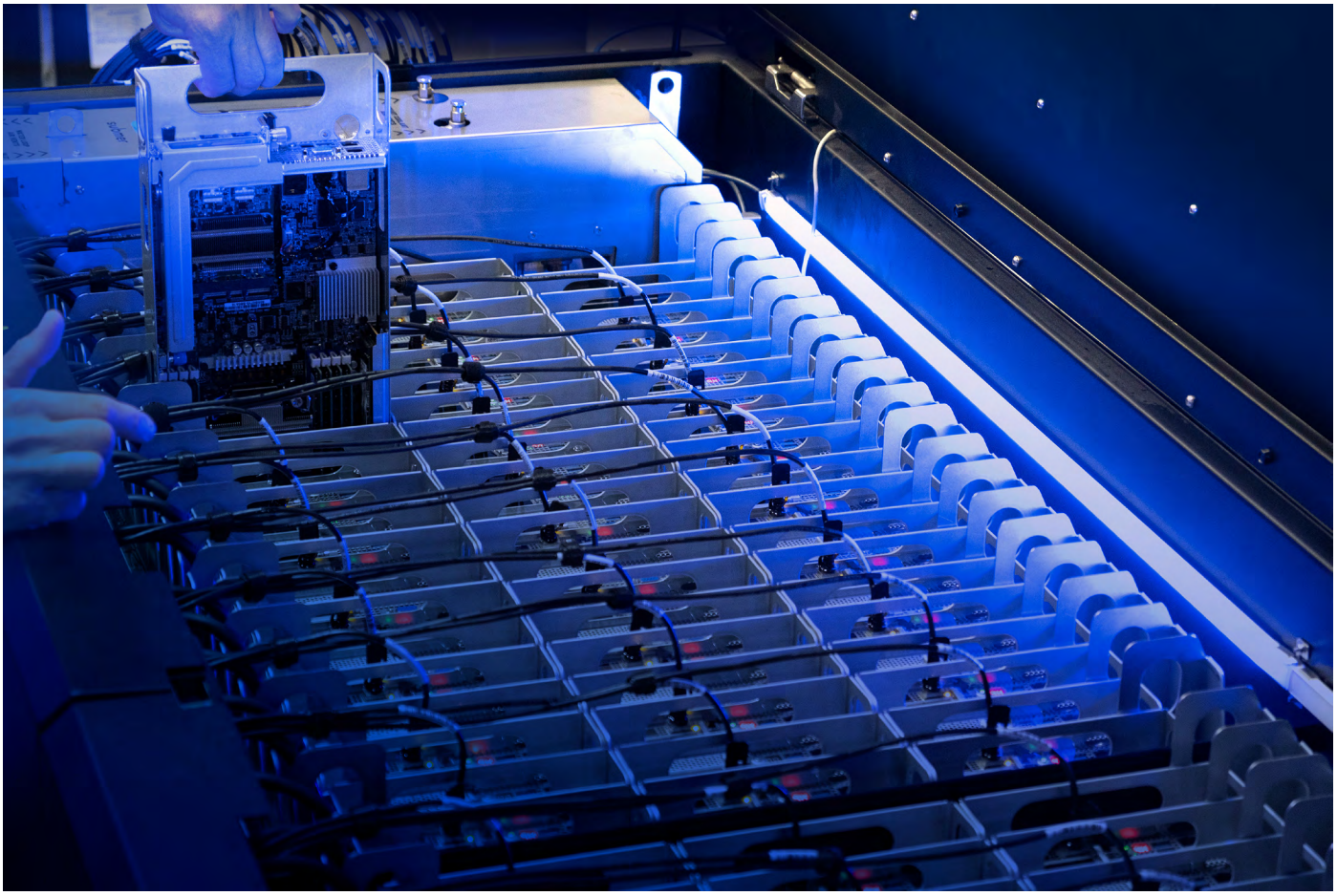
immersion cooling systems are now fully compatible with dielectric fluids.

- **Servicing and Maintenance:** Some believe immersion cooling systems are difficult to service and maintain. These systems are designed for easy servicing, with components accessible without draining the tank. This reduces downtime and simplifies maintenance operations.

- **Adaptability:** Immersion cooling systems are highly adaptable and can be customized to meet the specific needs of different data center environments. Whether it is a small-scale deployment or a large data center, immersion cooling can be tailored to fit the facility's requirements.

- **Flammability:** Concerns about the flammability of the fluids used in immersion cooling are unfounded. The dielectric fluids used in these systems are non-flammable, ensuring the safety of the data center environment.

- **Space Saving.** Immersion cooling, often perceived as requiring more space due to its horizontal configuration compared to traditional vertical racks, can actually reduce footprint by as much as 60%. Immersion cooling optimizes floor space utilization by streamlining cooling efficiency and eliminating the need for extensive air-cooling systems. This allows data center operators to reclaim valuable real estate, which can be allocated to other mission-critical infrastructure, enhancing overall data center density and operational efficiency.



Section 3 – Updated Designs and Applications

LIQUID IMMERSION COOLING DESIGN CONSIDERATIONS

As the adoption of immersion cooling continues to grow, it is essential to consider the design aspects of these systems to maximize their effectiveness. Several key factors must be taken into account when designing an immersion cooling system:

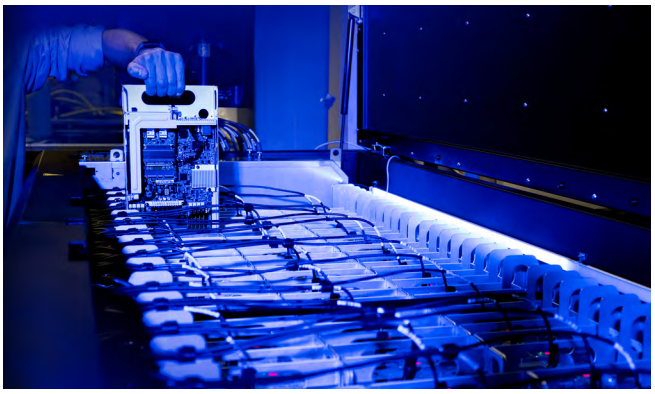
- **Direct-to-Chip Cooling:** Direct-to-chip cooling involves applying a coolant directly to the heat-generating components, such as CPUs and GPUs. This method provides more efficient heat dissipation than traditional air cooling, allowing for higher densities and improved performance.

- **Cold Plates and Immersion Options:** Cold plates are used to target specific hot spots within the hardware, while immersion options include both single-phase and two-phase cooling systems. Single-phase immersion involves submerging components in a liquid that remains in a single state, while two-phase immersion involves a liquid that boils and re-condenses to manage heat. Each option has its advantages, and the choice between them depends on the specific needs of the data center.

- **Dunking or Chassis Flooding:** These are different approaches to immersion cooling, each with its own benefits. Dunking involves submerging the entire server in a tank of dielectric fluid, while chassis flooding involves flooding the server chassis with fluid. The choice between these methods depends on the design and operational requirements of the data center.

UNDERSTANDING MODERN IMMERSION COOLING TECHNOLOGY

Immersion tank design is a critical aspect of the overall cooling system, as it directly impacts the efficiency and performance of the cooling process. For example, modern immersion tanks are designed with several key features that maximize cooling efficiency and ensure reliable operation.



- **Design Considerations in Hypertec:** Hypertec specifically engineered a chassis to help the liquid flow from bottom to top. With less obstruction, the liquid is more efficient at dissipating the heat, resulting in cooler temperatures. Further, a key feature to look for are purpose-built immersion machines. The optimized component layout of these servers simplifies maintenance, maximizes thermal efficiency, and improves reliability.

- **Retrofitted vs. Immersion-Born Servers:** Not all servers are equal regarding immersion cooling. Retrofitted servers are air-cooled servers that have been modified for use in immersion cooling systems.

  In retrofitted immersion-ready hardware, the liquid flow encounters obstacles as it navigates chassis housings and a backplane. Additionally, the absence of side openings, originally intended for air, further impedes the free flow.

  While they can be effective, they often do not perform as well as immersion-born servers designed from the ground up for immersion cooling. Immersion-born servers, such as those offered by Hypertec, provide better performance, reliability, and cooling efficiency, making them the preferred choice for modern data centers.

- **Optimization Within the Tank:** Effective immersion cooling goes beyond just submerging servers in the fluid. Careful consideration of fluid dynamics, cable management, and component placement is required to ensure optimal cooling performance. Proper cable management, for example, is essential to prevent cables from obstructing fluid flow and reducing cooling efficiency. Similarly, strategic component placement ensures heat is dissipated effectively throughout the tank.

SPECIAL SECTION: IMMERSION-BORN GPU SERVERS

GPUs are driving AI, so GPU-driven AI systems will be the next logical piece of componentry to be leveraged in immersion cooling. Hypertec's commitment to innovation is exemplified in its Immersion-Born GPU Servers, particularly the Trident iG610R-G6 and iG615R-G6 models. These servers are specifically engineered for immersion cooling, making them a powerful choice for handling the most demanding workloads in modern data centers. For these servers, there are several market-impacting use cases. The applications include, but are not limited to:

- **Classical Machine Learning:** Train models like decision trees, support vector machines, and clustering algorithms on large datasets. PCIe-based GPUs can efficiently handle these tasks and benefit from immersion cooling for sustained performance.

- **Moderate-Scale Deep Learning:** Develop convolutional neural networks (CNNs) for image classification or recurrent neural networks (RNNs) for sequence modeling that do not require the highest-end interconnects of SXM or OAM GPUs.

- **Edge AI Inference:** Deploy trained models for real-time inference in applications like object detection in surveillance systems or anomaly detection in manufacturing processes.

- **Batch Inference Processing:** This method handles large volumes of data that need to be processed through AI models, such as tagging images or documents, where latency is less critical but throughput is important.

Innovative Heatsink Design

One of the standout features of Hypertec's Immersion-Born GPU Servers is the advanced heatsink architecture. Unlike traditional air-cooled systems, these servers utilize a heatsink design optimized for the unique properties of dielectric fluids used in immersion cooling. The fins and overall heatsink structure are meticulously crafted to enhance heat dissipation, considering the fluid's viscosity to maximize cooling efficiency.

Hypertec has also leveraged forced convection add-ons for these heatsinks, further improving the heat dissipation generated by high-performance GPUs. As GPU thermal design power (TDP) continues to rise, with some GPUs reaching up to 1000W TDP, these innovations are critical. The design ensures that the servers maintain optimal temperatures even under extreme loads, something that traditional air cooling would struggle to achieve.

The Power of Immersion-Born GPU Servers

Hypertec's immersion-born GPU servers are not just about innovative cooling—they are about delivering raw power and reliability. Supporting the latest generation of PCIe GPUs, these servers are ideally suited for high-demand applications like machine learning, deep learning, HPC, and inferencing (inference training).

By being built from the ground up for immersion cooling, these servers eliminate the inefficiencies of retrofitted systems. They provide superior performance, greater reliability, and extended component lifespan while operating more energy-efficiently. For data centers looking to harness the full potential of GPU-driven workloads, Hypertec's Immersion-Born GPU Servers represent a robust and future-proof solution.

WHY IMMERSION-BORN IS BETTER THAN IMMERSION-READY [MINI-SECTION]

1. Optimized Design:

- **Immersion-Born:** Specifically designed for immersion cooling from the ground up, ensuring

all components are placed and configured to maximize cooling efficiency and performance.

- **Immersion-Ready:** These are retrofitted from air-cooled designs, which may not be fully optimized for immersion environments, leading to potential inefficiencies.

2. Improved Reliability:

- **Immersion-Born:** Reduced potential points of failure due to components being purpose-built for constant submersion in dielectric fluid.
- **Immersion-Ready:** Components may not be fully compatible or as durable when exposed to immersion cooling, leading to a higher likelihood of issues over time.

3. Enhanced Cooling Efficiency

- **Immersion-Born:** Features like vertical component placement and optimized fluid dynamics are integrated into the design, ensuring better heat dissipation.
- **Immersion-Ready:** Less efficient at cooling due to design compromises inherent in retrofitting air-cooled systems.

4. Longer Lifespan:

- **Immersion-Born:** Components are specifically chosen and tested for longevity in an immersion cooling environment, reducing wear and tear.
- **Immersion-Ready:** Retrofitted components might suffer from reduced lifespan due to incompatibilities with the cooling medium.

5. Lower Maintenance:

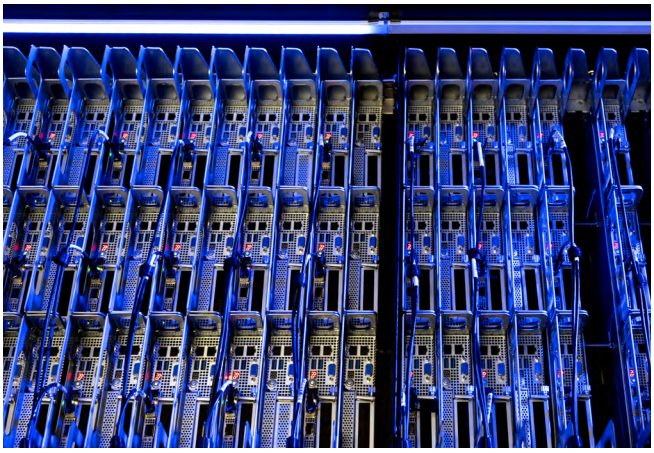
- **Immersion-Born:** Designed for easy serviceability within an immersion system, reducing downtime and simplifying maintenance procedures.
- **Immersion-Ready:** May require more frequent maintenance and adjustments due to the non-optimized nature of the retrofitted components.

KEY DESIGN CONSIDERATIONS

When implementing an immersion cooling system, it is important to consider several key factors to ensure the system meets the needs of the data center:

- **Vendor Considerations:** Selecting the right vendor is critical to the success of an immersion cooling implementation. Data centers should evaluate vendors based on their experience, product features, and support capabilities. Hypertec, for example, offers a range of immersion-born servers specifically designed for immersion cooling, providing a reliable and efficient solution for data centers. These include immersion-born GPU, compute, and storage servers.
- **Review Process:** Regular reviews and updates to the immersion cooling system are necessary to ensure it continues to meet the evolving needs of the data center. This includes evaluating the system's performance, identifying areas for improvement, and making necessary adjustments to maintain optimal cooling efficiency.
- **Challenging Vendors on Features and Capabilities:** It is important to challenge vendors on the features and capabilities of their products to ensure they meet the specific needs of the data center. This includes evaluating the scalability of the system, its ability to handle different workloads, and its long-term sustainability.
- **Serviceability and Maintenance:** Easy access to components and simple maintenance processes are critical for minimizing downtime and ensuring the longevity of the immersion cooling system. Hypertec's immersion-born servers, for example, are designed for easy serviceability, with tool-less PCIe card cages and other features that simplify maintenance operations.

Section 4 – Final Thoughts and Getting Started on a Liquid Immersion Cooling Journey



Before you jump on the liquid immersion cooling bandwagon, it's essential to understand your use cases and where liquid immersion cooling can make a powerful impact. Remember, depending on the component and the solution, a liquid can remove only a portion of the heat. That said, not all equipment will or should be liquid-cooled.

The interesting part is that, as mentioned earlier, liquid immersion cooling can create an all-encompassing solution where every component in your design is adequately cooled. So, designing around your use case is critical.

But how do you really know if you're immersion-ready? Check out this five-step checklist.

HOW TO BECOME IMMERSION-READY

1. **Understand Your Workload and Use Case:** The first step in implementing an immersion cooling system is understanding your data center's specific needs. This includes evaluating the types of workloads your data center supports, the computational power required, and the heat generated by these workloads. By understanding these factors, you can determine whether immersion cooling is the right solution for your data center.

2. **Calculate the Workloads, BTU, and Number of Nodes:** Once you clearly understand your workload requirements, the next step is to calculate your servers' heat output (BTU) and power consumption. This will help determine how many physical nodes can be accommodated in the immersion tank. For example, a 50U rack can hold up to 100 servers with a power draw of up to 150kW, depending on the cooling distribution units (CDUs) used.

3. **Architectural Review:** Before implementing an immersion cooling system, it is essential to conduct an architectural review of your data center. This includes evaluating the facility's physical layout, the infrastructure in place, and any modifications that may be necessary to support the new cooling system. Owning the building provides more flexibility in making these architectural changes, which can simplify the implementation process.

4. **Understand the TCO and Savings:** Implementing an immersion cooling system requires a significant investment, but the long-term benefits can outweigh the initial costs. To understand the full financial impact of the system, it is important to calculate the total cost of ownership (TCO), including both CAPEX and OPEX. Reducing energy consumption, lower cooling costs, and increased system longevity can lead to significant savings over time.

5. **Select the Right Partner:** Working with an experienced and reliable partner is critical to the success of an immersion cooling implementation. The right partner can provide the expertise and support needed to ensure a smooth and successful implementation. Hypertec, for example, offers a range of immersion-born

servers and comprehensive support services, making them an ideal partner for data centers looking to adopt immersion cooling.

Finally, it's less of a question of which is better and more of how to use liquid and air effectively. Liquid and air cooling will not be mutually exclusive when designing a good solution. This means working with the right partners to design your ecosystem will be critical. When it comes to liquid immersion cooling, this is a technology you should learn more about, as long as you have the proper use case.

RIGHT QUESTIONS TO ASK

As you begin the process of implementing an immersion cooling system, it is important to ask the right questions to ensure the system meets your needs:

- What are the specific cooling requirements of my workloads?
- How scalable is the immersion cooling solution?
- What are the long-term benefits of energy savings and environmental impact?
- How will the solution integrate with my existing infrastructure?
- What are the maintenance and operational requirements?

WORKING WITH IMMERSION COOLING LEADERS

As data centers increasingly adopt immersion cooling to address the challenges of modern workloads, selecting the right partner is crucial. Hypertec has established itself as a leader in this field, offering 360° solutions that are not only innovative but also specifically designed to meet the unique needs of today's data centers.

- **How Hypertec Immersion Cooling is Different.** What sets Hypertec apart is its focus on immersion-born systems—servers specifically designed for immersion cooling environments. We mentioned this earlier and discussed specific benefits between immersion-ready and immersion-born. These systems feature a unique architecture that enhances cooling efficiency, such as vertical component placement and strategic fluid dynamics, which are not found in retrofitted systems. This design not only improves heat dissipation but also extends the lifespan of the components, making Hypertec systems more durable and cost-effective over time.

- **Where They Create Direct Benefits.** Hypertec's systems provide direct benefits by significantly reducing cooling operational expenditures (OPEX) and allowing for higher server densities. Their immersion-born systems can achieve over a 5x increase in compute density, translating into more efficient use of space and energy within the data center. Additionally, Hypertec's systems contribute to a more sustainable and environmentally responsible operation by capturing and reusing heat.

- **Why Immersion-Born Systems are Powerful.** Immersion-born systems' power lies in their design, which is optimized for the specific challenges of immersion cooling. These systems eliminate the need for retrofitting, reducing potential points of failure and increasing overall efficiency. By leveraging immersion-born designs, data centers can achieve higher performance with lower energy consumption, all while maintaining the reliability needed for critical operations.

FINAL THOUGHTS AND A LOOK TO THE FUTURE

The demands on data centers will only increase as AI and HPC applications grow. Liquid immersion cooling, particularly immersion cooling, is no longer a luxury but a necessity for data centers looking to stay competitive and meet their customers' needs. Hypertec's immersion-born servers and advanced cooling technologies provide a reliable and efficient solution for data centers transitioning to liquid immersion cooling.

By adopting immersion cooling, data centers can achieve higher densities, lower energy consumption, and improved system performance, all while reducing their environmental impact. As the industry evolves, adopting liquid immersion cooling will be critical to maintaining performance, reliability, and sustainability of data centers.

The data center landscape will continue to evolve rapidly, driven by the increasing demands of AI and HPC applications. Adopting liquid immersion cooling will play a pivotal role in enabling data centers to meet these performance and environmental sustainability demands. Hypertec's commitment to innovation ensures that its solutions remain at the forefront of this evolution, providing the tools and support needed to lead in the new era of data center operations.

As the industry moves towards more efficient and sustainable solutions, immersion cooling will become more critical. Data centers that embrace this technology will be better positioned to meet future challenges and continue to deliver the high levels of performance and reliability that their customers expect.