MADISON AIR

THE WATER-ENERGY NEXUS IN THE AGE OF AI:

How Data Centers Can Optimize Water Use with Indirect Evaporative Liquid Cooling

AS data centers confront growing challenges in sustainability and resource use, across the United States and around the world, a clear-eyed view of water use is still the missing ingredient in delivering on an array of operational objectives—from efficiency and sustainability, to operational excellence and business continuity, to financial performance.

Most data center designers, owners, and operators pay close attention to water consumption at point of use, limited only to the water used within the four walls of the data center. However, they often miss the water-energy nexus where they can gain a deeper understanding of their full environmental impact—and, crucially, come up with the best opportunities to control it.

The bottom line is that data centers need water, whether it's used to cool chips or to generate the power that energy-hungry data centers require. With data centers facing demands for measurable results on environmental impact, and water scarcity on the rise, efficient use of both resources—water and energy—has long been an operational necessity. Now, as artificial intelligence (AI) drives up energy and thermal density for data center operations, more power-efficient liquid cooling is emerging as the only practical option.

A more rigorous approach to resource efficiency, in turn, delivers on the cost performance that end users expect. When data centers optimize their water and energy use, they increase the available electricity for their IT load by decreasing the amount needed to power ancillary systems such as cooling. That makes a sharp focus on water and energy an essential step in maximizing the cost-effectiveness of a data center's day-to-day operations.

THE SCOPE OF THE CHALLENGE

Data centers' electricity consumption has grown exponentially with the rise of more processing-intensive functions like video streaming, the Internet of Things, cloud computing, machine learning, smart and connected infrastructure, and now autonomous vehicles and AI.

Small wonder, then, that Power Usage Effectiveness (PUE)—the ratio of electricity a data center consumes to the power it delivers to computing equipment—has been stagnant for years, after falling dramatically over the dozen or so years ending in 2013. Andy Lawrence, Executive Director of Research at Uptime Institute, says data center operators aim to get their PUE ratios down as close as possible to 1.0, but the industry average has mostly been stuck between 1.55 and 1.67 for more than a decade—except when direct liquid cooling has brought the ratio down to 1.2 to 1.3, or better.

For some time, by shifting to more efficient systems, managing capacity better, and monitoring temperature and cooling more closely, data center operators were able to reduce operating costs in tandem with environmental footprints. More recently, however, any changes in PUE have been incremental and dependent on geography, with the ratio for centers using air-cooled chilling worsening modestly from 1.55 to 1.58 between 2022 and 2023. With the potential that AI will drive PUE up in many data centers, at least over the short term, it's time for new strategies and technologies to deliver the results that center operators and their customers expect—especially because power demands are certain to drastically increase, making each and every efficiency point more impactful.

The stakes are high. In 2018, according to a May, 2021 paper in the journal Environmental Research Letters, data centers around the world consumed 205 terawatt-hours (TWh) of electricity. Some 30% of data center servers were located in the United States, more than in any other country. "Previous studies found power densities per floor area of traditional data centers almost 15 to 100 times as large as those of typical commercial buildings," the paper stated, "and data center power density has increased with the proliferation of compute-intensive workloads."

By 2022, according to the International Energy Agency, data centers' electricity consumption stood at 240 to 340 TWh, accounting for 1% to 1.3% of global electricity demand and 1% of energy-related greenhouse gas emissions. In the United States, Newmark Group, Inc. expects data center power consumption to nearly double between 2022 and 2030, to a total of 35 gigawatts, with AI driving much of the increase. Legislators in both the United States and the European Union have begun to take note, and some data center operators are even considering options as extreme (and costly) as building their own onsite power plants to avert concerns about grid reliability.

And there's no sign that demand for data center services will stop growing: in mid-April, financial giant JPMorgan Chase announced plans to place

By 2022, according to the International Energy Agency, data centers' electricity consumption stood at 240 to 340 TWh, accounting for 1% to 1.3% of global electricity demand and 1% of energy-related greenhouse gas emissions

75% of its data and 70% of its applications in the cloud by the end of this year, after spending $2 billion on four new private cloud data centers in 2021 alone.

The May, 2021 paper in Environmental Research Letters addressed water consumption as well as energy, calculating direct use of cooling water as well as the volumes needed for electricity generation, including power consumption by water and wastewater utilities. The researchers estimated direct, onsite water consumption at 1.8 cubic metres per megawatt-hour, a considerable loss of efficiency compared to technologies like StatePoint. The indirect water consumption, almost entirely due to electricity generation, is nearly three times higher, producing a total of 7.1 cu.m./MWh—all compared to 0.9 cu.m./MWh for StatePoint.

That adds up to an annual operational water footprint for U.S. data centers of 513 million cubic meters in 2018, at a time when "many of the watersheds in the Western U.S. exhibit high levels of water stress, which is exacerbated by data centers' direct and indirect water demands."

Across the United States, the study found that one-fifth of data center servers depend on watersheds that are moderately or highly stressed, and nearly half of them rely at least partly on power plants in water-stressed regions. It recommended strategies for balancing data centers' environmental footprint and potential vulnerability by addressing water and energy in tandem.

FUTUREPROOFING FOR DATA CENTERS

The rising stress on energy and water systems, and the connections between them, make futureproofing an immediate priority for data centers to ensure business continuity and cost-effective operations, particularly in an era of accelerating climate change. The Center for Water-Efficiency at University of California-Davis states: **Every step of the water cycle—producing, moving, treating, and heating water, then collecting and treating wastewater—consumes energy. Just as producing water requires a lot of energy, harnessing energy requires lots of water.**

But concern about the water-energy nexus is not new. The U.S. Department of Energy [pdf] recognized the interdependence between water and energy systems in 2014. The IEA weighed in several years later with the observation that "low-carbon doesn't necessarily mean low water," calling for an "integrated approach" to energy development and emission reductions that also reduces water withdrawals.

These considerations are becoming more prominent with a focus on Environmental, Social and Governance (ESG) performance in business moving from the margins to the mainstream, across the U.S. and beyond, even as some customers de-emphasize ESG for the AI ramp. Bangalore, India-based ESG consultant Dr. Rakesh Varma writes that ESG: **…is as versatile as duct tape and has extended its reach across industries, and the data center industry is no exception. Bluntly put, neglecting ESG comes with a price tag—think decreased investor interest, tarnished reputation, and reduced employee morale.**

That general statement gets very specific, very fast when data centers begin thinking of ESG as an early warning for possible physical threats to routine operations. In its January, 2024 Drought Report, the U.S. National Oceanic and Atmospheric Administration (NOAA) cautioned that 23.5% of the contiguous United States (excluding Alaska, Hawai'i, and Puerto Rico, etc.) was in moderate to exceptional drought, with the most serious impacts falling in parts of the Southwest and the Southeast. The U.S. Center for Climate and Energy Solutions explains how: **climate change increases the odds of worsening drought in many parts of the United States and the world. Regions such as the U.S. Southwest, where droughts are expected to get more frequent, intense, and longer lasting, are at particular risk.**

Public expectations, political momentum, and the business case to address these risks have all been mounting since the adoption of the 2015 Paris climate agreement. Now, federal legislation like the U.S. Inflation Reduction Act has set off a global cleantech investment arms race, and the drive to decarbonize got another major boost from the final declaration of the December, 2023 COP28 climate summit. The consensus statement of 195 countries is already being cited in major policy and investment decisions, creating a new context and added urgency for data centers to consider energy- and water-saving innovations that are also essential to their bottom lines.

IT'S ALWAYS ABOUT WATER

It may not always be obvious when you first walk into a data center, or when a member of the public focuses in on a facility's direct water footprint. But it's still a reality: data centers' onsite energy and water use are joined at the hip. To conserve one, you almost always have to use more of the other.

> **The use of electricity and the use of water are inversely related. Having an elegant solution that is able to optimize for either power or water is one of our best benefits.**
>
> *— Marc Hayes, Director of Business Development, Nortek Data Center Cooling*

That means the most energy- and water-efficient answer for any data center depends on its location, its power sources, and the ability to vary its mode of use to meet different operating conditions. The full water footprint of electricity depends on how it's generated and how far it has to be transmitted, but it's always the case that water consumed at point of use is just a small fraction of the amount used for power generation. That means the best route to managing a data center's energy and water use is with a system that optimizes both.

Delivering on that promise is challenging for an industry that prefers to rely on known, familiar technologies that can meet the highest standard for reliability. That's a good reason to consider incremental, stepwise improvements that factor in the profound changes that data centers can expect to go through over a 30-year horizon.

Over that time span, as severe weather conditions become more common and water becomes less readily available for both power generation and water supply, there's a strong likelihood that the very definition of reliability will shift. In a less predictable physical environment, data centers will need to operate efficiently by relying more on local access to water or electricity, whichever is more readily available, while factoring in regulatory constraints and ESG considerations. A system that can adapt by changing the sequence of operations, rather than changing out equipment, will give centers the flexibility to adapt and be flexible to future demands.

NOW IT'S ABOUT AI

Video streaming, the Internet of Things, cloud computing, machine learning, and smart and connected infrastructure have all had an impact on data centers' water and energy footprint. The boom in artificial intelligence is the latest new dimension that strengthens the case for responsible, effective and efficient use of evaporative liquid cooling.

In today's operating environment, an air-cooled data center with 100,000 square feet of white space consumes 100 megawatts of electricity. With AI, the same space would require 500 MW due to the increasing electrical density AI chipsets draw along with the additional energy required to cool the much denser heat load. This increasing thermal density is also pushing the industry to a direct liquid cooling system design instead of the legacy air cooling systems used in most data centers today. Water is 24 times more effective at transferring heat than air, with the ability to carry 3,500 times more heat capacity than air using the same volumetric flow rate.

It all adds up to a simple new reality: In this new era in computing, air lacks the thermal potential to cool a more advanced generation of chips. And once again, even if it could, that five-fold increase in power demand has its own impact on upstream water consumption and may also drive up greenhouse gas emissions, depending on the local and regional energy mix.

But change won't happen overnight. A liquid-to-liquid heat exchanger such as a coolant distribution unit (CDU), direct-to-chip cold plates, and the manifolding needed to deliver the water to the servers constitute a major infrastructure design change that requires a significant investment from data center operators—and their first questions will quite rightly be about reliability and cost. The interim solution for many users, and the complete answer for some, is to use a combination of air and direct-to-chip liquid cooling that both rely on existing legacy building cooling systems. It's a blended approach that will minimize the infrastructure change needed to implement AI, though it will decrease the amount that can be added and reduce overall system efficiency..

Water is 24 times more effective at transferring heat than air, with the ability to carry 3,500 times more heat capacity than air using the same volumetric flow rate.

MADISON AIR

THE RIGHT FORM OF EVAPORATIVE COOLING: A BALANCED SOLUTION

Particularly in the age of AI, balanced cooling is the solution that delivers the real, measurable results that will satisfy investors and ESG representatives alike, pushing beyond onsite water consumption to optimize a data center's life cycle performance.

No two data centers are exactly alike, and their operational requirements are bound to vary over time. That's why owners and operators should be able to rely on a system like StatePoint that is built specifically for data centers and produces chilled water with a free cooling capable cooling water plant, delivered in a compact modular package. The system can adapt to a wide range of site-specific design criteria and be optimized for water consumption, power consumption, or both, simply by adjusting the sequence of operations. StatePoint promises a futureproof water cooling system with superior total cost of ownership, a durable and resilient system that minimizes the true environmental impact a data center's cooling system.

Already, conversions from older chiller/cooling tower and conventional direct evaporative cooling technologies to indirect evaporative liquid cooling are helping data centers reduce water intensity. Many hyperscale and colocation operators are moving toward these balanced systems, after adding up the huge amounts of water they consume with direct spray and other traditional evaporative cooling methods.

One of the world's leading hyperscale data center providers depends on StatePoint, and their experience shows that nothing succeeds like success.

For the first deployment, StatePoint enabled the company to reduce power consumption by 30% and water consumption by 90% while completely eliminating the need for mechanical cooling.

It's a very simple design, easy to maintain. There is no refrigerant that you need to worry about. It's just water going in and water going out.

— Jay Park, Chief Development Officer, Digital Edge

While this location was in a cooler temperature zone, water quality prevented a high cycle of concentration for an evaporative system, which made it paramount to save water at this site.

For the second deployment, StatePoint reduced total power consumption by 20% and water consumption by 20% in a very hot, humid location. Even in a harsh operating environment that made economization difficult, StatePoint allowed the company to reduce chiller plant capacity by over 50% and minimize total cost of ownership.

The system uses a patented, easy-to-maintain heat transfer membrane between the chilled water loop of a data center and the outdoor air, preventing cross-contamination between the streams. Not only does StatePoint provide more options for data center locations in the U.S.—in 2021, data center operators and the Government of Singapore announced a $17 million research program to support sustainable data center development in tropical climates, a move that would not have been possible without indirect evaporative liquid cooling technology such as StatePoint.

A SIMPLE, FLEXIBLE SOLUTION

At first glance, the StatePoint system might look like it demands a steep learning curve for users. But it mostly consists of familiar components like pumps, fans, and a chilled water coil. The membrane material and the controls are proprietary, but the system as a whole fits neatly into a routine HVAC maintenance schedule that successfully alleviates industry concerns about water management.

With more than 500 units deployed, more than 300 MW of critical IT load already being cooled, and more than 100 patents awarded or pending, StatePoint's success shows that a simple innovation like balanced cooling can deliver breakthrough results.

> **To me, we're setting a new standard for the industry...(It's) the best cooling system in the data center industry today.**
>
> — Jay Park, Chief Development Officer, Digital Edge

At Madison Air, our groundbreaking data center cooling systems forge a sustainable, efficient pathway to innovation, redefining industry standards and enabling our clients to revolutionize the world. With Madison Air's disruptive technology and expertise, we extend our commitment to excellence, improving reliability and performance across sectors and unlocking unparalleled business potential. Together, we aim to unlock the powerful potential of sustainable cooling for unprecedented success.

www.madisonair.com

By Jeff Hriljac, Sr. Director of Engineering
Nortek Data Center Cooling