COOLING GUIDANCE FOR DATA CENTER DENSITY, EFFICIENCY AND ECONOMY OF SCALE

Fluidity in business and global economies requires flexibility.

Eric Jensen, Vice President & General Manager, Data Aire
Dennis Cronin, CEO, DCIRN Americas
Muhammad Naveed Saeed, VP – Global Service Management, Uptime Institute
Amy al-Katib, Editor-in-chief, *Mission Critical*

Precise by Design

COOLING GUIDANCE FOR DATA CENTER DENSITY, EFFICIENCY AND ECONOMY OF SCALE

Fluidity in business and global economies requires flexibility.

Data Aire, Inc.

EXECUTIVE SUMMARY

In all cooling applications where significant change can be expected for a building's usage, future-ready flexibility is a foremost concern. Whether in regard to an office complex with movable walls or a colocation data center facility with movable customers, higher density, and variable data flow — in a world of constant digital transformation, having the capacity for growth is unquestionable. Demand and dependency for remote connectivity and data center resources grew exponentially during the recent pandemic, and building out the digital infrastructure quickly took priority.

Mission Critical magazine, in partnership with Data Aire, explores the major challenges facing design and facility managers regarding data center cooling. Data centers continue to pack more computing power into smaller spaces to consolidate workloads and accommodate processing-intensive applications, such as manufacturing AI, increasingly advanced consumer analyses, monitoring, and data handling. As a result, each rack consumes more energy and generates more heat, which puts more pressure on cooling systems.[1] In short, higher densities per rack require more cooling.

Several key scalability issues are addressed in this Guidance, including airflow management, water usage, sustainability, infrastructure, and other challenges for data center owners, designers, and operators. Information in this Guidance was collected from interviews and published works of Eric Jensen, vice president and general manager of Data Aire Inc.; Dennis Cronin, CEO of DCIRN for the Americas region; Muhammad Naveed Saeed, vice president of Uptime Institute; and Amy Al-Katib, editor-in-chief of Mission Critical, along with external research as footnoted.

ABOUT DATA AIRE

More than half a century of meeting and surpassing demanding expectations by every measure has made Data Aire a leading developer of innovative precision air control equipment and intelligent energy management solutions. With a dual mandate to protect assets as well as the environment, today Data Aire serves sensitive locations around the globe.

World-class manufacturing capability translates innovative designs into solutions— assembled with precision and built to last. Data Aire Management Systems include 5S methodology to optimize workplace performance and safety, and ensure ETL Intertek, UL Labels, and Intertek LAB certified standards. Data Aire is ISO 9001 certified and maintains its own AireLab™, which generates varied psychrometric conditions in support of product research and development and witness testing, prior to installation and commissioning.

dataaire
Precise by Design

IMPORTANT FACTORS FOR SCALABILITY

Scalability and what it means varies depending on the type of facility, according to Jensen. Data center site selection is based on multiple factors: proximity to business demand, network access and latency to users, geographical risk factors, and energy cost and availability. Building an infrastructure that can support the latest computer racks both being used today and that can be transitioned for the future is no small task. While many organizations have converted to colos and cloud computing, there are still many small- to mid-sized data centers and "server rooms" in operation. Many firms still see their data center cooling systems pushed to, and sometimes beyond, their limits.

"If you think in terms of the old-school telco hub and spoke kind of architecture, that same concept still applies in terms of scalability," Jensen said. "So, the internet giants who build their hyperscale facilities in largely remote areas — scalability means one thing to them there, and it means something different for them in their edge data centers.

"If you think in terms of the old-school telco hub and spoke kind of architecture, that same concept still applies in terms of scalability."

For example, in a greenfield site, in a very rural location, scalability is probably more of a function of repeating the same thing over and over again with the main question being is there enough space? However, when it comes to edge deployments, scalability might mean something more along the lines of load matching up and down as things change.

The initial data center size will often determine whether CRACs or CRAHs are used. CRACs may have insufficient lift to reach condensing units in multistory buildings, and refrigerant piping upsizing can sometimes be cost-prohibitive. CRAHs use chilled water and are not limited by distance or lift. However, water quality considerations are growing in importance as many data center structures are being built in areas susceptible to droughts.

Data Aire worked with one customer in a primary Southeast market and provided an ideal solution for scalable cooling. The client was building two facilities in succession with plans to build a third in the future and needed to ensure the anchor tenant could secure additional adjacent space as needed or move to a larger space in the new facility without disruption.

Data Aire provided the colo customer with cooling equipment that could scale from 25% to 100% load capacity based on workload. This allowed the customer to reserve additional space for the anchor tenant without overprovisioning.

In all cases, from small server rooms to hyperscale facilities, a modular thought-process for future data storage can take years to effectively design. Speed and flexibility, brought on by a recasting of the workforce due to the COVID pandemic, has reduced a typical two-year build-out for a data center by 50% to 70%.

dataaire
Precise by Design

AIRFLOW MANAGEMENT

The primary data center cooling technology used (air-cooled) is relatively mature, and, based upon cost-effectiveness at the hyperscale level, is thought to give ground in some marketplaces to liquid cooling. Nonetheless, air-cooled IT equipment (ITE) still dominates the majority of traditional data centers. Mainstream data center owners and users are risk-averse and don't accept change quickly, especially given the fact that most ITE racks are not typically fully loaded thus meaning the average power density does not yet justify the investment in more expensive technology. The exceptions will develop as hyperscale data centers with gigawatt-level campuses become the new normal.

Jensen pointed out that it's been about 20 years since the dot-com era, meaning there's a ton of existing aged space.

"There's a lot of replacement and refresh that's happening right now in the small to medium markets, and the first thing that everybody really ought to be doing when they're dealing with a project like that is to take a comprehensive site survey," he said. "Just a simple test and balance with a qualified technician service provider is a great way of understanding what you're dealing with as your current state before you before you embark on a refresh of any kind. A lot of people are missing that crucial step and they encounter problems as a result."

Rising energy costs already represent a significant part of operating expense. This will increase as densities are only likely to grow.

For air-cooled systems, such as DX or chilled water, either air or water provides heat transfer, and air is circulated across or around the equipment and controlled to varying degrees of complexity, i.e., hot/cold aisle containment, dedicated cooling, or dedicated airflow.

When it comes to air- or water-cooled technologies, there's not much difference as it relates to airflow management.

"If there is an existing central plant of any kind, then there are efficiencies to be gained through water-cooled technology," Jensen said. "But, depending on the size of the facility, if you're kind of in the mid-range or approaching large, it can create a single point of failure depending on how you built the chiller infrastructure."

Thermal management will continue to be a marriage of the potential variety of air-cooled and water-cooled technologies. Challenges for air cooling are increased densities and heavy processing load. Rising energy costs already represent a significant part of operating expense. This will increase as densities are only likely to grow.

In one particular instance, Data Aire worked with a client that was running a research and development (R&D) server room using air-side economization.

"For a long time, air-side economization was a big no-no in the industry, but it can be done right with proper planning and execution," Jensen said.

dataaire
Precise by Design

This particular R&D client experienced issues from the outset during the shoulder seasons. According to Jensen, air-side economizers can be hugely beneficial, but they require a very high level of commissioning to ensure the sequence of operations is correct. Frequently, designers are designing for the extremes, but it's the transitions that can become tricky.

In this case, it wasn't, so experts at Data Aire developed a sequence of operations between the operators and the cooling equipment, ultimately solving the issue.

"For a long time, air-side economization was a big no-no in the industry, but it can be done right with proper planning and execution."

DESIGN CONSIDERATIONS

Design decisions are not always easy and straightforward. There are thousands of SKUs offered by manufacturers for product categories, such as plenum and ducted returns, hot- and cold-aisle containment, in-row cooling, rear-door heat exchangers, and free cooling, to name only a few. Additionally, common design decisions include location, PUE, future rates, connectivity, network speed, power redundancy, green power options, cooling redundancies, and more.

There are a number of factors that influence a designer's choices:

- **Water Availability** — Fresh water is a scarce resource, and its availability is limited in some geographies more than others. Authorities in many countries are limiting and sometimes even disconnecting water supply to data centers.
- **Operational Complexity** — Water-cooled systems are relatively complex to operate compared to air-cooled systems and require trained staff to closely monitor and perform maintenance for continuous operation.
- **Ambient Conditions** — Higher ambient relative humidity levels combined with higher ambient temperatures can result in high wet bulb (WB) temperatures, a basic environmental condition for selection and performance of water-cooled technologies. The environmental conditions can limit the use and effectiveness of water-cooled technologies in certain geographical locations.
- **Sustainability** —Though water-cooled technologies provide lower PUE ratings over air-cooled technologies, higher water consumption rates pose a serious challenge to sustainability efforts. For water-cooled technologies, PUE must be viewed in concert with water usage effectiveness (WUE)
- **Economization/Indirect Free Cooling** — Water-cooled technologies are relatively more flexible to incorporate economization and indirect free cooling when environmental conditions are suitable.

dataaire
Precise by Design

SUSTAINABILITY AND WUE

Though PUE is the most recognized facility efficiency metric, it is only a part of an overall sustainability discussion. Still, renewable energy generation will absorb most of that discussion. The two primary renewable energy sources, wind and solar, are limited by variable and intermittent power output. Greenhouse gas emissions are another key factor in the sustainability equation. Carbon usage effectiveness (CUE) and WUE are additional metrics. The two are intertwined due to the tradeoff caused by the use of evaporative cooling to reduce cooling energy (as well as lowering PUE). This lowers greenhouse gas emissions but increases the use of water. However, the CUE and WUE ratio varies widely by geographic location; therefore, the amount of CO2 emitted also varies.

MAJOR CONCERNS OF HIGHER-DENSITY RACKS

It may appear that some concerns with higher density are a simple multiplication of that which already exists.

The benefits of facility redundancy to increase availability are sometimes traded off for energy efficiency, as designers take advantage of load sharing, improved IT virtualization, or multisite geo-redundancy supported by high-bandwidth architecture.

The potential for a variable heat load to exceed the cooling system capacity does exist at any time, and it's even more likely during high ambient conditions. Remote monitoring and control (exacerbated by the COVID-19 pandemic) are more important than ever. Sensors mounted in cabinets are a must for early warnings, and a plan to shut down the least critical systems so that more critical servers can remain operational is a failsafe plan more often advised.

The biggest factor of data center energy efficiency falls around cooling. Many colocation facilities have already moved down to lower PUEs, and, in some cases, larger hyperscalers are near 1.1, so where do we go from here? A colo, in particular, though able to easily meet the ASHRAE requirement of 64° to 80.6°F, will find most clients are not happy with 76° air blowing into a rack area. A balancing act exists between higher-density conditions and acceptable indoor ambient conditions and sound levels.

Additionally, multiple rack power densities in the same white space can result in an imbalanced cooling load. For example, for a white space designed to cater to 7- to 8-kW average power density that is populated with 3- to 4-kW power density racks in one part of the white space and 9- to 10-kW power density racks in another will create an imbalanced cooling load on terminal units. Terminal units serving higher densities will be pushing harder to meet

dataaire
Precise by Design

the set point as compared to the terminal units serving the lower densities. If the terminal units are grouped together to work at the same fan speed, the terminal units serving lower-density racks will be running at the same higher fan speed as terminal units serving the higher-density racks.

HIGHER DENSITY AT SCALE

"You have to satisfy the need for scalability," Jensen said. "And, higher-density cooling loads are still achievable in the same kinds of traditional ways. Whether you're talking about chilled water for some facilities or DX solutions (refrigerant-based solutions), for other types of facilities, both can achieve these higher densities in the traditional perimeter cooling methodologies without the need for completely rethinking the way that you manage your data center and the load coming from those servers."

"You have to satisfy the need for scalability, and higher-density cooling loads are achievable in the same kinds of traditional ways."

Greater airflow delivery per ton of cooling is a viable option, which is important because every operator is in transition mode — not only their IT architecture, but also their power side and their cooling infrastructure.

Those transitions are more manageable when cooling systems are engineered to order.

For DX solutions, that's achievable from a "good, better, best" scenario. "Good" was acceptable back in the 2- to 4-kW per rack days, but now, variable speed technologies are out there, and they can scale all the way from 25 to 100 percent. That's thanks to multi-fan arrays that can modulate from 25% to 100% for the delivery to scale over the life of the buildout or to accommodate fluctuations in workload throughout the day, season, year, etc.

And, what's more is they do so efficiently.

Many systems designed at the facility level lean toward dual cooling, which affords the redundancy of the infrastructure and also introduces the opportunity for economization.

dataaire
Precise by Design