THE OFFICIAL JOURNAL OF AIRAH
OCTOBER 2016 · VOLUME 15.9
RRP $14.95

Ecolibrium

Cooling the cloud

AIRAH

PRINT POST APPROVAL NUMBER PP352532/00001

Cooling the cloud

Best practice in data centre HVAC design is constantly being challenged as IT equipment becomes more robust and tolerant to higher operating temperatures and humidity. **Sean McGowan** explores the issues with a panel of industry leaders, including A.G. Coombs manager of data centre and telecommunication facilities, Mark Toner; Engineering Commissioning Services director, Neil Caswell, M.AIRAH; Aurecon global electrical leader Peter Greaves; and AECOM technical director for building engineering, Martin Howland.

Ecolibrium: What impacts have recent changes to ASHRAE specifications had on contemporary HVAC data centre design?

Toner: The changes reflect technology trends. IT equipment manufacturers are now designing and manufacturing IT appliances – servers, storage and network devices – with greater operating temperature and humidity tolerances.

And while the life of a data centre facility may be 15 to 20 years, the life of IT appliances is more like 3 to 5 years. So as technology is renewed through lifecycle upgrades, adoption of current ASHRAE specifications enables these higher temperature and humidity bands to be implemented.

BEST-PRACTICE DESIGN ELEMENTS

- Energy efficiency – the solution should maximise the benefits of outdoor temperatures to cool the data centre, directly or indirectly.
- Scalability – the free-cooling solution must be modular and scalable.
- Reliability – the free-cooling solution must be capable of supporting full load even when outdoor temperatures and humidity are not favourable.
- Standardised infrastructure design and componentry, and optimisation of facility design, including a modular approach to design and installation.
- Containment – data centres must be laid out in hot and cold aisles, and air streams should be separated with containment to ensure maximum efficiency of cooling architecture.
- When using direct free-cooling, the air must be filtered appropriately to reduce risk of contamination of IT systems.
- High-density requirements (→10kW per rack) require supplemental cooling solutions to ensure hot spots are avoided.

> "Data centres currently account for over 2 per cent of global energy consumption"

The impact of ASHRAE changes on design include more use of free cooling – direct and indirect – using either air-side or water-side economisation and less use of traditional large-scale chiller systems and perimeter CRAC cooling. There is also a reduced requirement for raised-floor systems in large greenfield builds, instead using air-handling plant – side or roof-mounted – to flood cold aisles, while containing hot aisles and exhausting hot air straight outside.

It should be noted, though, that Class 1 and 2 parameters remain unchanged in recent changes to ASHRAE 2011. The ASHRAE 2008 version had the greater impact on changes to design with the introduction of free cooling. In our experience, some clients have readily adopted the changes while others have clung to more conservative, tighter tolerances on room conditions.

But the ASHRAE standard is not the only standard that is applied in data centre design in Australia. In particular, the American Telecommunications Industry Association (TIA) ANSI/TIA-942 – Telecommunications Infrastructure Standard for Data Centers is also applied.

Howland: The ASHRAE 2015 Thermal Guidelines provide an expansion of the allowable ranges of supply-air condition to the rack, for various classes of environment – in some cases up to 45°C!

Depending on the IT manufacturers present in the white space and the data centre manager's service level agreements (SLAs), this change allows for greater supply-air temperature ranges, and allows designers and operators to achieve an increase in annual hours where economisers (or free cooling) can be utilised.

Caswell: With the expansion of the allowable operating envelope for internal conditions, we have seen a move to incorporate free cooling strategies within HVAC systems. This has manifested itself in many different ways with varying success, dependent on implementation and local climate.

In addition, there is often a limiting factor for colocation data centres where the tenants insist on environmental conditions that aren't always compatible with current practice and relate to historic agreements. This can also provide additional challenges.

Greaves: More free cooling solutions are being considered for data centres as a result.

Air temperature that is supplied is in line with the ASHRAE guidelines (18–27°C). Supply-air temperatures of up to 27°C will need outside-air temperatures at 25°C or less in order to gain significant benefits from free cooling.

Cooling technology is not, of course, only concerned with controlling temperature. It is also important to control humidity to prevent condensation in the case of high humidity, and electrostatic charge problems in the case of low humidity.

Other considerations are server manufacturers warranting their equipment at the increased operating temperature, and the wellbeing of personnel working in the data centre.

Mark Toner

Neil Caswell, M.AIRAH

Peter Greaves

Martin Howland

> The main challenges in delivery are around a robust and thorough testing of all of the services

Ecolibrium: Are we seeing a new paradigm in best practice emerge?

Greaves: High performance has become expected.

In the past, successful data centres were those that achieved high levels of redundancy requirements. Today, however, those high levels of redundancy have become the expected performance norm.

In addition, increasing demands in performance go hand-in-hand with the need for increased amounts of energy and water consumption. In the past, the challenge was to design for these requirements whereas now, responsible development dictates we mitigate these requirements and taken a step further: design for resource scarcity in some areas.

Toner: From what we see, there is a definite trend to scalable and modular power/cooling systems, especially in the commercial sector where facilities are built out in stages to best align capital expenditure to client demand.

Free cooling is incorporated in all new commercial data centre builds, which is not surprising given cooling energy costs have a big impact on the operator's cost base and overall market competitiveness.

With modular scalable designs, prefabrication of plant rooms and cooling and power modules enable an accelerated construction program and faster deployment of new facilities. Time to market is becoming important in this new digital economy.

Howland: There is an emerging shift in focus to a "total system solution" rather than a discipline-by-discipline focus.

From site identification through to design deployment, that accounts for scalability, constructability, fast turn-around from site identification to full functionality. Delivering high reliability at low cost requires focus on coherence, simplicity, repeatability, symmetry and consistency in design approach across all disciplines.

Early intervention and collaboration with experience construction managers further enhances buildability and minimises construction time.

Caswell: There are certainly more variances to how the data hall is served – from all services viewpoints.

This can range from the plant usage such as CRACs (computer room air conditioning unit), IRCs (in-row cooling) and AHUs (air-handling unit), through to the method of handling airflow – hot aisle containment, chimneys, RFM delivery, and the provision of stand-by power from traditional battery UPS systems, flywheel-based UPS systems and alternative power distribution methods.

We are also seeing a lot more modular solutions for small and medium-sized data centres, though very large data centres still tend to utilise large central plant.

Ecolibrium: What are some of the key elements that make best-practice design?

Howland: We are seeing an increase in the uptake of purely modular data centre solutions, where clients and operators limit their initial capital outlay to the initial base infrastructure only, with further expenditure on subsequent future phases of IT infrastructure delayed until the demand requires it.

Our design focus is to achieve fast deployment of simple and reliable systems that meet a specific performance objective and cost plan. Our mantra is, "Simple, smart and easy to operate and maintain".

Caswell: One of the main elements is a push to higher chilled-water temperatures, allowing for more opportunity to use water-side economisers. The use of direct and indirect evaporative cooling technologies is also being implemented successfully, especially indirect evaporative cooling in more tropical applications.

The key message for any mission-critical facility is to keep the systems as resilient as possible. Any outage can have devastating effects. With this in mind, simplicity is key – though perversely the way to achieve simplicity can be complicated!

Toner: Hot and cold aisle layout combined with containment systems is now well understood and practised. Overlay this with a free- cooling architecture and the ability for supplemental cooling to accommodate any high-density technology racks, and you're well placed to provide a flexible, energy-efficient environment. In these critical facilities, all building services need a high level of coordination and integration. BIM (building information modelling) is best practice for achieving this during the design and construct phase.

Greaves: Besides reduced electricity costs, there are additional incentives to going green. Data centres currently account for over 2 per cent of global energy consumption, so companies can benefit greatly from positive media coverage if they are willing to invest in lowering their electricity demand from data centres.

Google, for example, has bought the entire electricity output of four as-yet-unbuilt wind farms in southern Sweden to power their data centre. Facebook, similarly, inaugurated a massive data centre in Lulea in Sweden in 2013, and this facility runs entirely on renewable energy generated by nearby hydroelectric schemes. This has gained both of these companies positive media attention.

To prove a data centre is green, it is now possible to grade a data centre against the LEED certification program. There hasn't been a large uptake in this yet, but its set to become more popular in the future, especially as interest from large banking institutions mounts.

Ecolibrium: What are some of the challenges, and are they common across the world or are some unique to Australia?

Caswell: The main challenges in delivery are around a robust and thorough testing of all of the services.

It isn't enough to just test components or systems individually, but it's paramount to verify that the systems operate in a holistic manner, with the electrical, mechanical and associated services supporting each other seamlessly.

FROM 2010 TO NOW

Mark Toner, manager of data centre and telecommunication facilities for A.G. Coombs, lifts the lid on how data centre energy efficiency has improved since 2010.

"Energy consumption is now being metered across ICT systems and critical data centre infrastructure systems – for example, cooling systems," says Toner.

"There is a much stronger understanding, even down to the rack PDU socket level and how much energy is being consumed. Better measurement equals better management ability."

Toner says an important approach to energy efficiency is to maximise the return-air temperatures through matching the cooling air flow to the data equipment needs, and minimising surplus and bypassed air with containment systems.

"This allows free-cooling systems the best possible temperature differences, and therefore the best opportunity for higher utilisation."

According to Toner, other initiatives that have improved efficiency in data centre systems include:

- Free cooling (direct and indirect)
- Adoption of hot and cold aisles
- Variable speed fans in CRACs and AHU plant
- More efficient lighting systems
- More efficient UPS systems
- Modular approach, standardisation of electrical and mechanical streams/capacities for optimal performance of each stream/module.
- Greater scalability in key systems including cooling plant and UPS plant enabling systems to operate at utilisation levels where they are most efficient.

"In the ICT layer," he says, "virtualisation has been the key to optimising server and storage utilisation, and consolidating from many to fewer physical devices."

> Data centre design is not about inventing systems, but about putting available systems together to provide the correct solutions

A typical Integrated System Test script, which is used to record the results of a simulated failure across all services, can run into as many as 100 pages of verified points or actions and take a number of hours to fully test. On some of our projects, to complete all of the necessary integrated testing has taken weeks of sustained effort.

Greaves: Ten years ago, we were designing data centres for an average electrical load of 500kW/m². But with increased technology, IT demands and smartphones, this load has continued to escalate and we can be designing for 1200kW/m².

We are faced with the dilemma of having higher pollution and carbon loading in the industry, when overall there is genuine commitment to sustainable development. This translates to cooling efficiency being a serious issue.

Howland: One of the key challenges in unlocking improvements in data centre efficiency around the world is getting data centre managers (DCMs) to allow higher supply-air temperatures to be supplied to the rack. Most commonly, DCMs are limited in going to thermal conditions beyond the ASHRAE recommended limits – and usually much tighter – and are often limited by their contractual SLAs.

This is driven by limitations in IT equipment, with many data centres hosting legacy kit with reliability issues in conditions outside the standard SLAs. For new-build data centres, there is greater scope for clients to push these limits with their IT suppliers, to mandate that only equipment that is tested and rated for higher temperatures will be deployed to the data centre, thereby allowing increases in SLAs and design conditions and efficiency.

Toner: One of the primary challenges is to find ways of economically achieving low PUEs. This could be seen as a local issue, in that the size of the more common data centres in Australia is relatively small, and this presents its own challenges in achieving economies of scale.

Ecolibrium: How important is the PUE ratio in best-practice design? Is it the be-all and end-all when it comes to measuring data centre energy efficiency, and what's an acceptable number?

Greaves: Yes. Simply put, the common metric for assessing the energy efficiency of a data centre is PUE.

We are seeing considerable variation in efficiency depending on the technology being employed. Legacy data centres can use 1kW of cooling for every 1kW of IT load, leading to PUE ranging from 1.8 to 2.0.

Free-air-cooling technologies such as indirect evaporative cooling are achieving PUE values of between 1.15, and 1.28 and the incentive to develop alternatives is inhibited by the reality of diminishing returns. Current good practice is considered to be a PUE of 1.2 using free cooling in temperate climates.

Howland: Economical, annualised PUEs for Sydney or Melbourne will sit between 1.2 and 1.4.

While PUE is still the most common terminology used to define data centre efficiency, it is not the be-all and end-all efficiency rating from our perspective.

A number of different ratings provide users with different aspects of efficiency. Data Centre Infrastructure Efficiency (DCiE) provides users with an efficiency metric more suitable for use by the IT users, whereas Carbon Usage Effectiveness (CUE) takes into account the embedded carbon content in the energy sources being consumed.

Caswell: PUE is often written into the SLAs of a data centre, and there are cost penalties for not achieving it. However, resilience is more important.

The PUE can also be the win-lose criteria in the process of engaging tenants for many colocation data centre sites, especially considering that many PUEs are now stipulated at varying loads rather than at the traditional 100 per cent load point.

Toner: PUE is always important, of course; however, not at the expense of reliability and availability. In mission-critical facility design, energy efficiency considerations are always on the basis of, "What can we do without compromising the robustness of this system?"

PUE is a metric that everyone understands and has a pretty transparent formula sitting behind it.

It should be noted, however, that while it provides you with a relative measure of your data centre's energy efficiency, it can provide an illusion that may distort actual energy consumption trends. PUE does not necessarily provide you with a game plan for driving down net consumption.

In view of this, Green Grid is launching a new matrix to better assess data centre performance. This promises to address both data centre performance and resilience.

Ecolibrium: Is free cooling now a standard part of contemporary data centre HVAC design? Where do such opportunities exist and how is it realised in the data centre environment?

Caswell: Free cooling is becoming more widespread as an initiative within HVAC design, though its implementation varies. Design solutions may be water or air-based dependent on the desired targeted area. There is certainly a push to incorporate free cooling where possible to drive down PUE.

One of the primary challenges when using outside air for free cooling is with the stable pressurisation of the data halls when modulating between controlled settings. We are currently seeing indirect cooling solutions being more prevalent than direct solutions.

Toner: Free cooling is certainly a worthy contributor and can be used in most of Australia's regions. The challenge is to ensure that the payback justifies the plant investment, which can quite large for air-side free cooling.

Free cooling is best suited to new facilities, where the designer has a blank canvas and where a solution can be designed to maximise the free-cooling potential of the target geographic location.

For existing facilities, free cooling is often limited in its application due to the current design of the legacy data centre facility. Consequently it is a difficult and/or expensive challenge to overcome in converting old systems.

So yes – for greenfield locations in supporting climate zones, it is.

Greaves: Within Australia, I would say its popularity has risen dramatically in the last three years.

Free-cooling opportunities are possible in many locations, especially if the air temperature that is supplied is in line with the ASHRAE guidelines (18-27°C). Supply-air temperatures of up to 27°C will need outside-air temperatures at 25°C or less in order to gain significant benefits from free cooling.

Data centre managers then need to decide whether they are going to use direct or indirect free cooling. I tend to prefer indirect via a heat wheel or heat exchanger so that outside-air contaminants or humidity levels do not restrict the use of free cooling. This type of indirect free cooling is particularly suitable for areas where the temperature falls below 19°C and the humidity is below 60RH for more than 2,500 hours per year.

Howland: Free cooling is most definitely a big part of contemporary data centre HVAC design, with opportunities in the design of new centres in accordance with the allowable ASHRAE environmental conditions.

These elevated operating conditions allow mechanical equipment to operate at lower capacities and with lower input energy. Opportunities are not limited by system type, with different methods applicable to systems, such as air-side or water-side methodologies.

Ecolibrium: Is cogeneration or trigeneration being used in data centres, and does it lead to lower PUEs?

Toner: Cogeneration or trigeneration is being used to a much limited extent compared to previous times. It really has gone off the boil post the carbon tax – gas costs have escalated and operational and maintenance costs are also high, and electricity costs remain relatively low. The ROI – return on investment – is often difficult to justify.

Its effect on PUE is not an assured improvement, and there are sometimes technical challenges around provision of minimum base load in new data centre facilities.

It is worthwhile pointing out that the other green benefits offered by cogeneration or trigeneration, including a cleaner, lower energy source and lower water consumption, are still worthy.

Greaves: While cogeneration is used in data centres, it is not widely used and lower PUEs are not always achieved. It would need to be designed and operated correctly.

We tend to see more rotating machines utilised as DRUPS – diesel rotary uninterruptable power supply – than cogeneration.

Howland: Cogeneration does provide an improved PUE, although often is sensitive to tariffs and challenged with maintenance procedures and costs. Most often, the decision to go with cogeneration or trigeneration is aligned with a corporate desire to achieve carbon reduction more so than life-cycle costs.

Caswell: We have seen little evidence of cogeneration use in data centres. Issues around maintenance, getting gas to site, and the reluctance to rely on absorption chillers make its use cost- prohibitive to most operators.

Ecolibrium: Does the privacy and security around data centres limit the amount of information sharing going on in our industry?

Caswell: Up to a point, yes. However, given the limited number of engineers within the industry who have recent experience within the data centre market, their knowledge is often shared through employment changes within the industry.

Data centre design is not about inventing systems, but about putting available systems together to provide the correct solutions.

Greaves: I spend a lot of time in data centres, and I find people very keen to collaborate and share experiences. ▮