HOW COOLING HAS HEATED UP THE DATA CENTER DISCUSSION

FOCUS blogger and managing director at Operational Intelligence, Robert Tozer

It’s been interesting to look at how the data center industry has developed in the past five years. When it comes to cooling, much of the advice has remained the same, although there have been changes in what has been adopted by operators. Changing the focus from reliability to energy optimization does not happen quickly in a risk-adverse industry, though there is now a greater understanding of the common ground of reliability and efficiency and a growing body of specialized knowledge that stimulates innovative solutions to address these challenges.

Increasing policy and cost pressures have forced the industry to look at its energy efficiency and improve the way data centers are cooled — usually the largest energy consumer after the IT load in a legacy facility.

In line with Gartner’s Hype Cycle, the key technologies for achieving highly efficient data centers require:
- Good air management/physical containment
- Variable speed drives on CRAH/CRAC fans
- Free cooling
- Higher air supply temperatures from CRAH/CRAC units

There is now a greater understanding that we need to control the temperature of the air at the server inlet, rather than the ‘room’ or CRAC return temperature. Air containment inside the data hall is normal for new facilities and has been retrofitted into many existing sites in order to better manage cooling resources. ASHRAE TC 9.9 has worked with IT hardware vendors to expand the recommended and allowable environmental ranges so that data center cooling systems can run at warmer temperatures and capitalize on less refrigeration and more free cooling.

This has allowed low-PUE designs to become a reality, although more often than not the measured performance does not meet the design targets. Even with containment in place, there is normally room for improvement — with closer analysis, systems often underperform when energy improvement measures have been implemented in a ‘tick box’ manner, rather than with a clear understanding of the objectives and impact. For example, many contained systems still have a large degree of common bypass and control issues. Metrics on cold and hot air segregation are useful tools to help operational teams analyse performance.

A modular and scalable approach is now commonplace, but low-energy designs usually have plenty of scope for further optimization once in operation, particularly at low loads.

Most new designs now incorporate some type of free cooling and the indirect air system is starting to gain popularity — this avoids the problems of outdoor air contamination and humidity control, allows an increased number of free cooling hours in most climates, higher feasibility of zero refrigeration and results in CAPEX reduction on refrigeration and electrical infrastructure. As IT exhaust temperatures increase, rack exhaust containment systems are likely to become more popular as they separate operations teams from the areas of hot air.

In future, as the Cloud grows and IT utilization increases, we can expect to see the trend for higher densities continuing, along with higher server delta Ts and exhaust temperatures. This will make good air management even more important and increased numbers of stakeholders will require improved communication for overall technical ownership. The higher densities might drive us away from air cooling and towards direct liquid cooling of hardware, or a combination of both methods.

As the industry matured, the way it thinks about data center cooling has changed. However, many operators still have some way to go before they understand and prioritize cooling best practices enough to benefit from the operational cost savings, particularly the ones with good returns on investment.