Data Centers: Overcooled, Overcooked, or Just Right

The Thermal Monitoring and Analysis Route to Efficient Energy Usage Within the Data Center White Space

Introduction

Industry and customer requirements are driving data center operators toward implementing environmental monitoring capabilities. Environmental monitoring across the data center provides data, analysis, and actionable intelligence that allows operators to drive up efficiency and reduce energy consumption, which generates savings. For instance, modular thermal efficiency solutions which use temperature and pressure to control the data center white space environment provide clear targeting criteria to achieve operator goals and can offer cost neutral systems as well as massive on-going savings.

For this paper, data center white space refers to the server, switch, and data storage area, which also house the air handling equipment to those cabinets. The other areas within the data center, which are just as essential and house the chillers, uninterruptible power supply (UPS), and general operational areas, are gray spaces. This discussion focuses on understanding monitoring and controlling both gray and white areas, which is necessary for a range of standards compliance and customer requirements.

There are significant differences in implementation between data centers focused on an intelligent monitoring strategy and data centers that sustain a 'basic' deployment. However, the cost and environmental benefits for using an intelligent monitoring system, especially within modular asset and connectivity management solutions, provide an engaging argument (Figure 1).

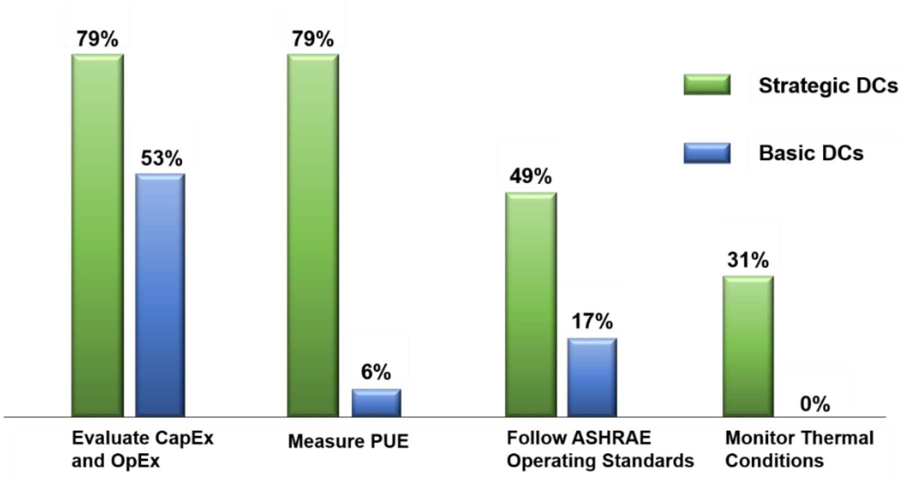

*Figure 1. Reasons Data Center Owners Invest in Environmental Monitoring.**

*Source IBM Estimates

The cost of outages and their apparent frequency are rising (Figure 2). Recent outages at a Wikipedia data center shut down the site for hours while Microsoft's Outlook email service shut down for 16 hours due to an unplanned outage. Both incidents were caused by data center servers overheating and the servers' automated systems shutting down. In each case, there was no monitoring and control system in place to avoid these unplanned outages.

Cost comparison for partial unplanned outages and complete unplanned outages. All three studies show complete outages are more than twice as expensive as partial outages.

Cost for partial and total shutdown
Comparison of 2010, 2013 and 2016 results

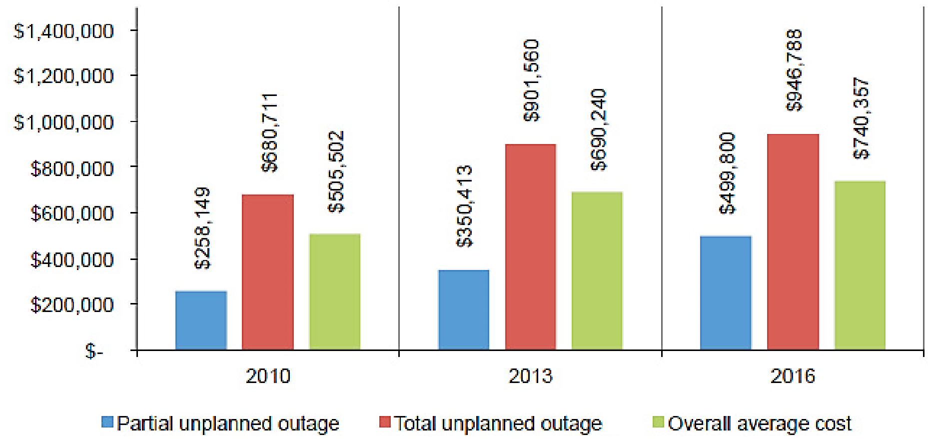

Figure 2. Cost of Partial and Total Shutdown.*

International standards, including ASHRAE TC 9.9, ETSI EN 300 and EN 50600-2-3, provide a framework for best practice, compliance, and determining suitable Information Technology Environments (ITE). The technology is available for operators to design systems and policies around guidelines for data processing environments, which can stop these overheating incidents, reducing, or eliminating user disruption and possible financial penalties.

Airflow Management and Cooling System Control - ASHRAE TC 9.9 Guidelines

ASHRAE TC 9.9 guidelines recommend new thermal design and installation practices for airflow management and cooling system control in a data center. These guidelines help data center operators avoid overheating and loss of networking equipment functionality. Below is a sampling of the guidelines that are relevant to this discussion.

1. Cabinet/Rack level – Instrument and monitor the inlet temperature and RH for racks and cabinets at the bottom, middle and top of the cabinet, maintaining a specified recommended (18-27°C) as well as allowable (15-32°C) thermal ranges
2. Containment level – in addition to 1. – With a cold aisle containment system, the hot aisle temperature can be in the range up to 50°C; instrument and monitor the outlet temperature at the top of the rack and cabinet. When using hot aisle containment system, then temperatures across the room must be monitored
3. Data Hall level – in addition to 1. and, or 2. – Humidity and temperature needs to be monitored near each CRAC/CRAH at the Supply and Return. Relative humidity is recommended at 60% RH and allowable at 20% to 80% RH

*Source Ponemon Institute

4. Airflow Management & Cooling System Control – Airflow management and cooling-system control strategy should be implemented. With good airflow management, server temperature rise can be up to 20°C; with inlet temperature of 40°C the hot aisle could be 60°C

Air pressure management is also an essential component in a robust and effective airflow management and cooling system within the data center white space.

White Space Then and Now

Data centers are energy intensive. Operating white spaces with hundreds or thousands of servers uses vast amounts of energy and generates significant heat, which must be addressed. It is not unusual for the cooling system of a facility to use as much, or possibly more, energy than the white space it supports. Today, a well-designed white space with a monitored and controllable cooling system may use a greatly reduced level of energy. In many cases, the latest developments in thermal planning, monitoring, and cooling optimization are saving hundreds of thousands of pounds, euros, and dollars in energy costs, as well as pre-empting problems and providing a more resilient and reliable data center.

To facilitate this important change, a mind-shift is needed. The previous pervasive concept was to create a cool environment where hot equipment (servers and switches) would have cold air passed across the live surface and the hot exhaust drawn away. This heating, ventilation, and air conditioning (HVAC) solution required a great deal of energy to reduce the 'air inlet' temperature, to that needed to lower the temperature across the hot equipment, while the hot exhaust is often expelled and wasted.

Today's white space processing equipment has higher operating temperatures; therefore, this allows the data center industry to develop alternative cooling methodologies which take advantage of intelligent environments. The warmer the white space operational temperature, the less energy is needed to equalize the 'air inlet' temperature. Device inlet temperatures between 18-27°C and 20-80% relative humidity (RH) will usually meet the manufacturer's operational criteria. Gaining importance is the capability to monitor and control the recommended environmental range, including temperature and RH and to maintain an allowable environmental envelope, where the systems are operating at optimum performance.

Operating within a higher temperature environment means that the high-performance computing (HPC) servers are working closer to their maximum operational parameters. If, for example, a massive spike in processor activity takes place, and many more servers are brought online to cope with the capacity, while at the same time a generator fails and the UPS back up is not 100% efficient, this could lead to cooler fans not coming online quickly enough and the servers overheating and shutting down. As discussed above, an unplanned outage can cost the data center revenue in terms of customer compensation, damaged reputation, and future customer contracts.

A requirement to secure optimal performance is the capability to intelligently monitor the white space thermal environment and analyze the data generated in real time to provide actionable intelligence to maintain effective white space operations.

Deploying an Environmental Monitoring and Cooling Optimization Solution

There are three distinct levels that data center operators need to consider and implement for an environmental monitoring and cooling optimization solution.

1. Monitoring - alarms and notification - ASHRAE provides guidelines for sensor distribution within the ITE white space. The latest wireless sensors offer thermal, thermal with humidity, and pressure nodes. These are easily configured on a wireless mesh network and allow for simple, fast, and secure device deployment, offering a highly resilient self-healing, scalable, and efficient sensor network.
2. Cooling Optimization – airflow remediation and floor balancing. This includes designing the airflow metrics and using computational fluid dynamics (CFDs) to allow modeling of environmental scenarios in line with the operator's goals. It also includes using blanking plates, evaluating containment options, and installing perforated floor tiles to ensure optimized air pressure across the critical pathway and using real-time cooling mapping software to provide heat maps of the white space.
3. HVAC control and dynamically matching cooling to IT load. Real-time data analysis across the system allows for actionable intelligence. The system constantly analyzes data to improve airflow management to reduce energy use. Real-time control allows the system to dynamically maintain the optimized state through airflow management via fan speed adjustment, using pressure nodes for readings, and air temperature management via temperature set point adjustment using temperature monitoring.

Conclusion

The amount of energy needed for cooling the world's data centers is expected to triple over the next ten years, both in costs and environmental impact. Customers, environmental lobbies, and government will drive the industry to develop efficiencies in its energy use. Data center operators are now investigating and implementing intelligent systems to begin that pathway to continuous analysis, data assessment and dynamic optimization. Targeting specific data center operator goals for these latest environmental and cooling management systems will provide actionable intelligence that offer a greatly reduced ROI timescale, by using simple and efficient asset utilization.