

HOW AND WHY THERMAL OPTIMIZATION PAYS OFF

Cost-efficient cooling considerations combat challenges of shifting IT loads.

Today, there is a need for facility and IT managers to develop a different approach in how they address data center cooling challenges.

They will find themselves facing even more heat as digital growth and shifting IT loads place increasing demand on data centers. Such scenarios are caused by increasing computing power, resulting in a need to optimize cooling. In fact, the forecasted 20 billion connected devices by 2020₁ will have forever transformed data centers and their IT managers' involvement from their current state. These devices require the internet, computing resources, and applications to connect. This, in turn, increases IT loads and generates heat, putting more pressure on data center and IT managers to provide effective, economic cooling.

By Jay Hendrix

Jay Hendrix is the portfolio head of data centers, North America for Siemens. He is responsible for data center strategic alignment, solutions management, and market investment and strategy in the North American region. He has more than 25 years of industry experience in product management, controls installation, and wireless and technical integration. He can be reached at Jay.Hendrix@siemens.com.



Thermal optimization is one such approach which allows facility and IT managers to shift from traditional cooling to more innovative and holistic processes. It matches cooling to IT load at every critical spot in the white space and optimizes a data center's chilled water plant to save money. This article will identify some of the factors IT and facility managers may consider in thermal optimization to reduce energy costs, and improve efficiency and equipment life in their data centers through proper cooling.

WHY THERMAL OPTIMIZATION?

While there are many different types of data center cooling processes available today, there are few that provide a comprehensive view, assessment, and optimization of cooling systems (supply side) and the major heat producers — the IT server equipment (demand side). Blade servers started the heat revolution in data centers. In the data center, blade servers cut both ways. They require only a tiny footprint, which is great when office space costs \$40/sq ft. However, they generate enormous amounts of heat for their size. Since dozens of blades, each with thermal output, can occupy a small area, cooling airflow is a must.

Thermal optimization works well for both facility and IT managers. Facility managers benefit from improved equipment life through optimal operation, while IT staff may witness improved reliability and increased available capacity. Overall, thermal optimization typically results in a significant elimination of waste.

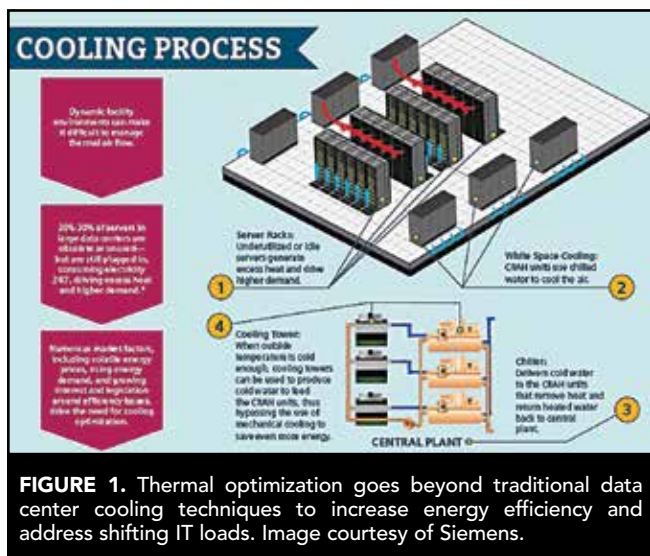


FIGURE 1. Thermal optimization goes beyond traditional data center cooling techniques to increase energy efficiency and address shifting IT loads. Image courtesy of Siemens.

It also provides impressive real-world results applicable to average applications or settings. Face it — the typical hospital or retail data center is not a Google or Facebook operation built for efficiency. The industry watchword for years has been constant uptime at any cost. While uptime is still crucial, there are other concerns when it comes to operations.

Cost matters. For example, power and cooling make up a large amount of a data center’s operating expense (OpEx). A realistic goal for an average data center is to keep cooling costs below 20% of the overall OpEx spend. As such, power consumption cannot and should not be ignored as a source of efficiency gains and cost savings. In addition, thermal optimization approaches can yield hundreds of thousands of annual savings with two to three year paybacks.

Efficiency metrics also matter. Power usage effectiveness (PUE), one of the most standard industry measures developed by The Green Grid, is a simple division of the total power load divided by the IT load. A perfect score would be 1.0. To measure the reality of how many industry data centers are close to this, let us take a look at some examples. Facebook’s Oregon facility, designed for super efficiency, has a sterling 1.06 PUE.² This may not be a realistic goal for the typical enterprise’s data center. An average data center’s PUE is in the range of 1.8 to 2.0. A 2.0 PUE means that half of a data center’s power spend is going to equipment, such as chillers, uninterruptible power supplies (UPS), and the like.

KEY THERMAL OPTIMIZATION CATEGORIES

Thermal optimization addresses three main categories of data center cooling challenges: reliability, consumption, and capacity. A comprehensive thermal optimization approach needs to address all three areas.

The PUE equation focuses on the efficiency of the infrastructure. However, unless you take a look at the server side you still may be cooling to an artificially high load. PUE does not look at this side. Thermal optimization examines both sides to balance supply and demand and eliminate waste.

The cooling systems in most data centers were designed based on peak IT heat loads. However, most data centers rarely operate at peak load. In many cases, a thermal optimization project simply needs to be implemented to realize dollar-saving benefits.

The immediate by-product of thermal optimization is a reduction in energy costs. Yet thermal optimization goes much further and can help an enterprise increase reliability, decrease consumption, and increase available capacity. The result is a win-win-win project.

ADDRESSING RELIABILITY CHALLENGES

White space cooling systems need to operate efficiently to provide optimal reliability to a data center. Ensuring that this happens often means addressing the low-hanging fruit in your data center — addressing those data center hotspots often caused by poor air or water circulation, excess airflow, and unevenly distributing cooling. Water and air distribution inefficiencies can be identified with data, and simple fixes like closing perforated tiles can reduce hotspots. However, IT and facility managers may find better use in a comprehensive thermal optimization approach because it can improve reliability and cut costs.

Analytics and automation can identify where workloads are running. Combined with thermal analytics, those workloads can be moved and adjusted to reduce risk of failure. This would naturally help avoid hot spots created by putting a workload where there is not enough air distribution capacity, and would allow an IT or facility manager to maximize the utilization of their infrastructure.

Managing workloads proactively can also be accomplished by identifying which influencers are causing hot spots or other issues in your data center. For example, temperatures at IT cabinets or server racks within a data center can influence where those hotspots are. Once these influences are known, they can be used to identify reliability risks, to help make provisioning decisions, and to optimize the air delivery and energy consumption of the cooling equipment. Having this level of information about the state and behavior of the airside is also important for optimizing the performance of the waterside, which is where the largest portion of the cooling energy is consumed in a chilled water cooling system.

ADDRESSING CONSUMPTION CHALLENGES

On the facility side, integration between airside equipment and building management system (BMS) is necessary for control and optimization. The same holds true for chilled water plant equipment. Most mechanical equipment comes equipped with integrated controls provided by the original equipment manufacturer (OEM). This equipment easily communicates to the BMS via standard protocols. In the physical plant, the chilled water and condenser water pumping systems, the chillers, and the cooling tower fans must be part of a thermal optimization approach.

Important diagnostic information and control values can be sent from the mechanical equipment to the BMS. This is necessary for white space cooling and chilled water plant control. By monitoring outdoor air temperature and other criteria, the BMS can determine the right time to use free cooling versus mechanical cooling.

Quick Quiz Questions

Q. What percentage of a data center's physical and virtual servers do no useful work at all and should be considered zombie servers?

- A. 50-80%
- B. 30-50%
- C. 10-30%
- D. Less than 10%

A. After looking at tens of thousands of instances in data centers across North America, Siemens found that 35% of servers are complete zombies. The correct answer is B, 30% to 50%.

Q. By the end of 2016, how many data center managers will have performed thermal retrofit upgrades to their systems or plan to do so?

- A. 25%
- B. 35%
- C. 50%
- D. None of the above

A. The correct answer is C, 50%. Are you one of them?

The BMS also integrates to the data center infrastructure management (DCIM) system to share important information for monitoring of capacity and reliability. IT and cooling load data can be sent to the DCIM system which in turn provides capacity information to IT staff. The BMS provides important alarm information too. This alarm information can be used to kick start predefined work flows within the DCIM system to address key service needs.

A successful thermal optimization approach will improve cooling distribution to server inlets to eliminate hot spots. It will reduce wear and tear on chillers and pumps through optimization. The key here is to run units in their efficient "sweet spots." Thermal optimization serves to identify problem areas where cooling needs cannot be met, and will alert IT systems of this sort of deficiency so loads can be shifted.

ADDRESSING CAPACITY CHALLENGES

Energy and deliverable capacity are interdependent. Often "conservation methods" reduce deliverable system capacity. Other times, what is done in the name of energy conservation results in a transfer of energy among these subsystems with no net savings realized — or, worse, an actual increase in energy requirements.

On the capacity side, a good thermal optimization approach will result in a reduction in energy use and will reduce server waste by identifying ghost or zombie servers. These servers provide no use and could be shut off or even decommissioned.

There are a couple of potential savings areas when zombie units are isolated. First, their power consumption could be reduced to zero if servers are decommissioned or turned off. Second, it may give the IT staff a reason to update the old servers to newer, more energy-efficient servers. This closes yet another

faucet that is dripping data center dollars.

All of this improvement and efficiency builds data center capacity without significant added expense. Countless operations have discovered that optimization of chiller equipment provides a decrease in energy cost and an increase in available capacity.

IDENTIFYING COST SAVINGS

As we identified earlier, there are several ways to save costs with a thermal optimization approach. However, it is important to know your key performance indicators (KPIs) before starting any project where thermal optimization will be applied. KPIs are the metrics used to measure your goals and are a proven way to drive business value. Once the goal is set, it is measured against the KPI. Your list of KPIs is dependent upon your key goals as an organization. In this context, KPIs typically fall into three overarching categories: managing system operations and compliance, optimizing performance and productivity, and protecting lifecycle investment.

One large financial institution faced inefficient cooling control in their central plant and white space areas. Their first remedial step was to implement chilled water plant optimization. Through the project the supply water temperature was increased, allowing the data center to use the existing plate and frame heat exchanger longer. As a result, they cut their utility bill by over \$200,000, and the utility provided about \$200,000 more in rebate support.

Next, they looked at the white space itself. A lack of adequate control and sensing capability throughout the data center required expensive and wasteful overcooling of the data halls. Following that step, the financial institution needed to utilize a white space cooling optimization solution. Wireless technology was used to monitor rack inlet temperatures. The wireless sensors coupled with optimized control allowed facility managers to raise inlet temperature to the ASHRAE-recommended (American Society of Heating, Refrigerating and Air-Conditioning Engineers) 80.6°F. The result was that over half of their computer room air handlers (CRAHs) shut down completely, reducing IT energy consumption. The return temperature to the chilled water plant increased, meaning the chillers operated more efficiently, resulting in less wear and tear and less energy wasted. The data center experienced a 71% drop in kWh, putting more than \$240,000 back on the balance sheet. On top of that, there was a utility rebate of \$150,000. Altogether, the business was able to accomplish return on investment (ROI) in less than two years.

NEXT STEPS

It is important to understand that data center cooling challenges are not insurmountable. Regardless of your data center's inefficiency levels, taking a thermal optimization approach can start small.

Typically, there are levels of evaluation and assessment. The basic level will require someone to visit the data center and take about a day in the typical corporate environment to gather information (name plate data, utility bills, etc.). It will define for IT, data center, and plant managers what sort of problems might exist and where improvements can be made.

Thermal Optimization Takeaways

- Thermal optimization addresses cooling supply and demand for a holistic approach.
- Reliability, consumption, and capacity are key factors in a thermal optimization approach.
- Facility improvements are no longer constrained by the lack of visibility into data centers.
- Opportunities to achieve added savings by lowering OPEX 30-40% can be quantified to justify thermal optimization investment with the right processes in place.
- Both facility and IT team leaders must be involved in discussions around thermal optimization to argue its case and achievable benefits.

That can be followed up by in-depth second and third tier evaluations that go deeper into the operation. Every day thermal optimization is postponed is another day when energy costs leak out of the budget.

With a typical payback period of two to three years, most implemented projects that use a thermal optimization approach can stand on their own. On top of that, across the country utilities recognize data centers are huge consumers of electricity. Many utility companies offer substantial rebates for taking steps to

remedy inefficiencies in the data center. Often, these incentives are in the six-figure area.

Call someone and take action with a thermal optimization approach. There are qualified companies that can do a good assessment for you to identify problem areas in the white space or with server loads. Do not stand by and watch others save money through better coordination of their data center cooling.

WORKS CITED

1. Gartner, Inc., Nov. 10, 2015, "Gartner Says 6.4 Billion Connected 'Things' Will Be in Use in 2016, Up 30 Percent From 2015:" <http://www.gartner.com/newsroom/id/3165317>
2. Oregon Live LLC, April 19, 2013, "Facebook offers peek at how its Prineville data center performs:" http://www.oregonlive.com/silicon-forest/index.ssf/2013/04/facebook_offers_peek_at_how_it.html ■

Read this article online at

www.missioncriticalmagazine.com/jayhendrix