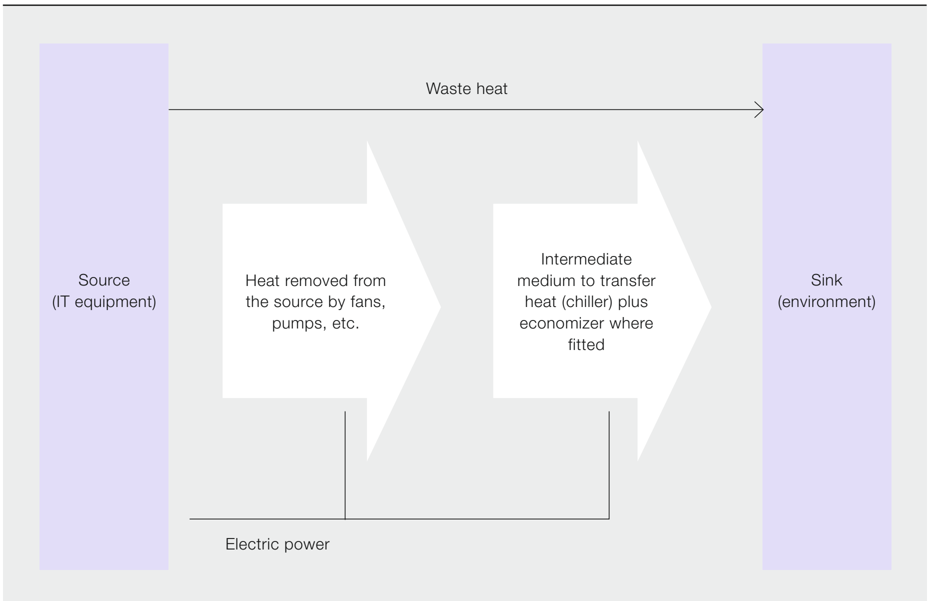
Keeping it cool

Optimal cooling systems design and management

SHRIKANT BHAT, CARSTEN FRANKE, LENNART MERKERT, NAVEEN BHUTANI –
Heat generation is a cause for major concern in data centers. Indeed, up to 45 percent of the total energy used in a data center can go to just cooling the server racks [1]. This figure is set to rise as servers become ever more compact and, as a result, power densities increase. Cooling technologies, power management and associated control systems are rapidly evolving to combat this escalating heat problem. A modern cooling system that can rise to the challenge of this situation must adopt a radical approach and focus on improved energy efficiency, integrated management and high reliability for the entire data center. ABB's experience in managing critical power systems and complex industrial processes stands it in good stead to take up the cooling challenges a data center presents.

Title picture
A large part of the energy consumed by a data center ends up as waste heat. Dealing with such a large heat load in such a small volume requires sophisticated cooling technology and techniques.
Photo courtesy: © 2013 Michelle Kiener

1 Heat flow in data centers



Until recently, heat management techniques in data centers were based on the methods used to cool buildings. Thermally, a server was treated as an "equivalent human" and this assumption worked fairly well. However, the heat flux from commercial microprocessors has increased from around 1 W/cm^2 to 100 W/cm^2 over the last decade and this is expected to rise further [2]. This represent a massive increase on the demands faced by any cooling system.

Cooling in data centers involves the transfer of heat generated from IT equipment (source) to the environment (sink) in a two-step process: The heat is first transported by a medium (air or liquid) out of the server racks and then it is rejected to the environment ➔ 1. Both these steps consume electrical energy. The target of cooling efficiency measures, then, is to reduce the energy required to remove the heat and recover and reuse as much of it as possible. This can be achieved through innovations in the design of the cooling system itself as well as by inventive operating strategies – eg, smart sensing and monitoring, and integrated system management.

Cooling system design and management has several important areas and it is worthwhile to examine each of these ➔ 2.

Novel cooling designs

There are various cooling technologies at different stages of commercial maturity and some of these show promising results ➔ 3. Aisle containment, for instance, is practiced commercially and can improve system efficiency by up to 30 percent [3]. On-chip cooling is at a preliminary research phase and has been reported to achieve cooling of up to 15 °C for heat fluxes as high as 1,300 W/cm^2 [4]. Liquid cooling is expected to reduce cooling energy consumption by as much as 50 percent compared with conventional air-cooled systems and is being commercialized now. Membrane air drying and evaporative cooling is reported to reduce energy requirements by up to 86.2 percent compared with conventional mechanical vapor compression systems [5].

The waste heat from a data center can be augmented by solar thermal energy to drive an absorption chiller, thus reducing power usage effectiveness to less than one (absorption chillers use the hot water from the primary cooling loop, and solar heat on occasion, to drive an additional chiller loop).

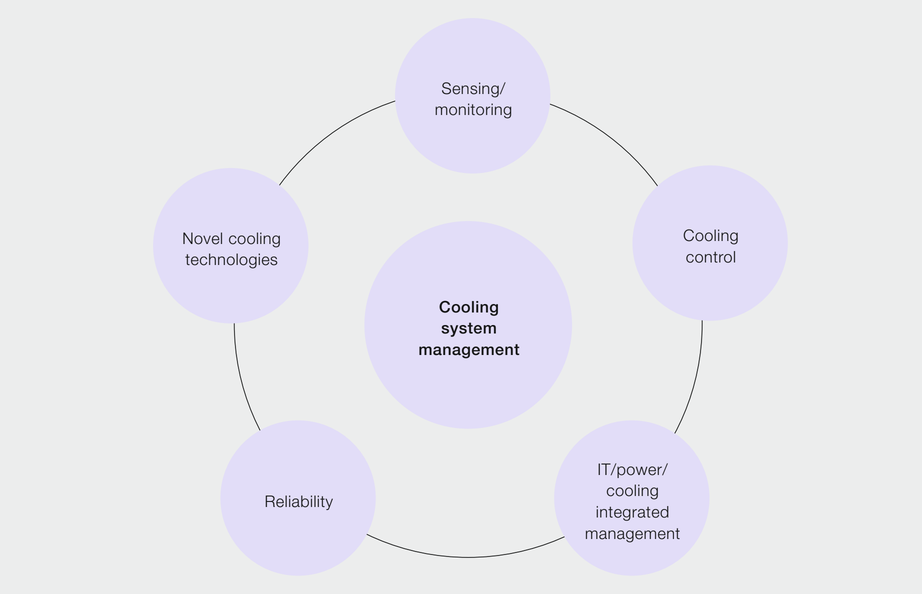
Liquid cooling, absorption cooling and evaporation-based cooling have already been practiced in other industries. However, data centers pose unique challenges in terms of the nonhomogeneous heat generation associated with highly dynamic load behavior and the requirements for high reliability. ABB has expertise in ensuring high reliability for critical power system components along with extensive experience in integrated process management. This capability can help address the challenges posed by integration of novel cooling technologies with data centers.

Monitoring and sensing

The first step in managing and controlling cooling is to monitor the thermal behavior of the data center. Hot spots are a major cause of concern and these can be detected using infrared sensing or wireless sensors. Soft sensors that combine data already available with detailed computational fluid dynamics models, or empirical models, are another important tool.

The clear target of cooling efficiency measures, then, is to reduce the energy required to remove the heat and recover and reuse as much of it as possible.

2 Focus areas for cooling system management



IT load management is often decoupled from the cooling and power systems – so IT jobs are started with no regard for the cooling or power required. To avoid this, coordination of all three subsystems is required.

It is also important to benchmark emerging technologies:
– What are the current cooling technologies and their limitations?
– What advanced solutions can be integrated with the cooling system?
– Up to what level is integration or adaptation feasible and what are the system limitations?
– What is the impact of a new solution on the reliability of the overall cooling and IT system?
– What will be the value (cost benefits, return on investment, etc.) of the newly added resource?

ABB has demonstrated the use of concepts such as infrared sensing, wireless communication, soft sensing and fingerprinting across different application areas in the power and automation domain. This know-how can be extended, with suitable adaptations, toward data center performance monitoring.

Cooling control

A data center cooling unit has a chiller, cooling tower, pumps and thermal storage → 4. It often also has an economizer, which provides a form of "free cooling." Economizers complement the existing cooling by drawing in colder outside air and using it to reduce chiller energy consumption. The external air passes through one or more sets of filters to catch particulates that might harm the hardware. It is also conditioned to an appropriate relative humidity.

Optimizing such a cooling system in an integrated way involves minimizing the net cost of power while ensuring that cooling requirements for a given IT load are met. This often results in a complex demand-response problem that involves inputs of weather forecast, energy prices and load-versus-efficiency curves for all the equipment involved. An integrated cooling approach involving only economizer integration, along with model predictive control strategies for temperature control, has been shown to reduce cooling management costs by up to 30 percent [6]. This situation could be further improved by the use of additional storage and demand-response management to exploit energy price variation.

A modular approach

Modular cooling units allow data centers to expand their capacities incrementally. So popular have such units become that they now constitute a de facto design standard. However, they present a challenge to integrated cooling control as there is an interaction between them and related common facilities such as the chiller, evaporator and economizer. This poses additional constraints on the integrated cooling control problem described above.

ABB's cpmPlus Energy Manager has the ability to handle such integrated demand response management problems to help customers realize additional benefits.

Optimizing a cooling system in an integrated way involves minimizing the net cost of power while ensuring that cooling IT requirements are met. This often results in a complex demand response problem.

3 Drivers for novel cooling design and representative cases

Driver	Representative cases	Comments
Thermodynamic efficiency	Aisle containment On-chip cooling	Efforts targeted toward minimizing energy and exergy loss by localized heat removal and avoidance of mixing different temperature streams.
Materials	Liquid cooling Membrane air drying and cooling	Novel materials are offering higher efficiency and more rapid heat removal.
Waste heat recovery	Absorption cooling	Cooling with waste heat recovered from data centers is one of the most promising options.
Renewable integration	Solar cooling	Solar cooling is one of the most promising options for using renewables for data center cooling.

Integrated management of power, IT and cooling

In almost all existing data centers the IT load management is not coupled with the cooling management or the power supply. That means the IT load management software makes an independent decision when to start new IT jobs, or when to migrate running jobs, without any consideration for the cooling or power required. This "selfish" behavior can reduce the power used by the IT equipment, but at the expense of a higher cooling energy consumption.

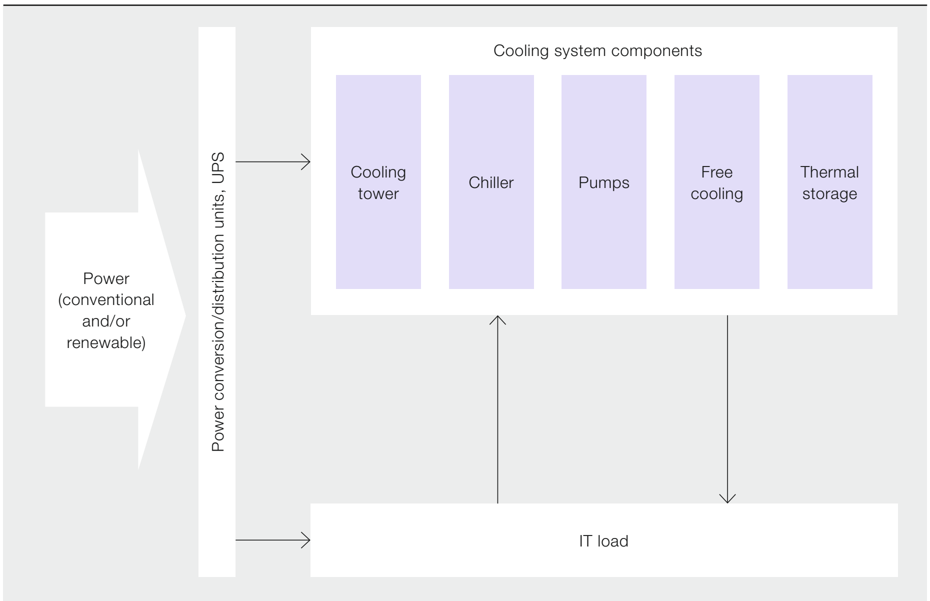
To avoid such scenarios, coordination of all three subsystems is required. Furthermore, it is also necessary to have a dynamic and predictive IT load management tool so that the data center location and corresponding time-varying energy provisioning can be taken into account. Such an advanced load management, which could be integrated with ABB's data center infrastructure management (DCIM) system, Decathlon™, can lead to energy savings of 20 to 40 percent [7].

Such a solution can also directly accommodate maintenance aspects – for example, by load sharing among servers and their related cooling devices to equalize component aging. It can help with power management too: Should cooling requirements or energy prices reach critical values, Decathlon, for example, can dynamically lower the voltage supplied to components or reduce the clock frequency to reduce cooling needs and energy consumption. This technique, called dynamic voltage and frequency scaling (DVFS), is performed in such a way so as to ensure IT jobs do not violate their given service level agreements (SLAs). IT jobs can be migrated to other servers, too, to save power or cooling. In the past, such migration was limited to a very few applications that supported check-pointing, but increased use of virtualized servers now makes migration easier.

Coordinated management can be extended to incorporate resources not just from one data center but from several, geographically distributed data centers. This can lead to further energy savings of 5 to 10 percent. The main advantage of spreading IT loads between data centers is that the installed capacity per data center can be smaller than the maximum that would be needed were the data center to operate in isolation, as some resources can be shared. This indirectly also leads to a better energy usage. An IT load-sharing strategy exploits time-of-day energy price variations and ambient temperature differences between locations. Energy price predictions and cooling forecasts can easily be extracted from Decathlon, leaving only the information flow to the IT load management to be established.

IT load management across data centers provides benefits but is also subject to legal and logistical constraints. For example, data may be bound to a certain jurisdiction, thus limiting migration options. In addition, security aspects and data pro-

4 Schematic of data center showing cooling, IT and power components



tection become important if the data centers belong to different legal entities. Furthermore, the additional energy demand and communication costs involved in migration must be considered.

Reliability

Fluctuating humidity, poor air quality and temperature variations are the main phenomena related to the use of an economizer that impact reliability. To improve reliability, the intake air quality can be monitored and if it drops below certain standards preventive actions can be taken. For example, Decathlon can automatically close external air intake vents and switch to another means of cooling when air quality standards are threatened.

Hot spots also detrimentally affect reliability. Effective monitoring and control can deal with these without overprovision of cooling for the entire data center. This directly reduces energy costs.

Another approach used to increase reliability is to regularly maintain or replace critical equipment before failure occurs. The intervention can occur after a fixed period defined by the mean time between failures or the manufacturer's warranty. However, a fixed period approach is not ideal. Load profiles, and environmental and operating conditions, might vary from the average values indicated by the manufacturer so it is better to tailor maintenance and replacement for each piece of equipment individually. A loss of performance or unusual equipment behavior can indicate upcoming failures, so monitoring the operating conditions of critical components can allow better planning of maintenance and replacement actions.

For example, the voltages across several capacitors of a power converter show massive voltage drops and unusual oscillations shortly before the power adapter fails. If such deviations are monitored and automatically tracked, preventive actions like repair or replacement can be initiated just when they are needed. Equipment downtime is thus decreased as failures are predicted and corrected before equipment drops out. Consequently, reliability and availability of the data center are increased. In addition, unnecessary maintenance and replacement costs are eliminated.

ABB has demonstrated its capability to monitor and ensure system reliability in a wide range of mission-critical applications in industrial power and automation settings. This experience puts ABB in a perfect position to manage mission-critical data centers for customers, especially when tools like Decathlon are available.

Advanced load management, which could be integrated with ABB's Decathlon DCIM system, can lead to energy savings of 20 to 40 percent.

Shrikant Bhat
Naveen Bhutani
ABB Corporate Research
Bangalore, India
shrikant.bhat@in.abb.com
naveen.bhutani@in.abb.com

Carsten Franke
ABB Corporate Research
Baden-Dättwil, Switzerland
carsten.franke@ch.abb.com

Lennart Merkert
ABB Corporate Research
Ladenburg, Germany
lennart.merkert@de.abb.com

References

[1] J. B. Marcinichen *et al.*, "A review of on-chip micro-evaporation: Experimental evaluation of liquid pumping and vapor compression driven cooling systems and control," Applied Energy, vol. 92, issue C, pp. 147–161, 2012.

[2] J. B. Marcinichen *et al.*, "On-chip two-phase cooling of data centers: Cooling system and energy recovery evaluation," Applied Thermal Engineering, vol. 41, pp. 36–51, 2012.

[3] Subzero Engineering Inc., (2013, August) Hot aisle containment. [Online]. Available: http://www.subzeroeng.com/containment/hot-aisle-containment

[4] C. Ihtesham *et al.*, "On-chip cooling by superlattice-based thin-film thermoelectrics," Nature Nanotechnology, Vol. 4, Issue 4, pp. 235–238, 2009.

[5] El-Dessouky *et al.*, "A novel air conditioning system: Membrane air drying and evaporative cooling," Trans. IChemE, Vol. 78, Part A, pp. 999–1009, 2000.

[6] R. Zhou *et al.*, "Optimization and control of cooling microgrids for data centers," HP Technical Report, 2012.

[7] W. Nebel *et al.*, "Untersuchung des Potentials von rechenzentrenübergreifendem Lastmanagement zur Reduzierung des Energieverbrauchs in der KIT," Technical report, OFFIS Institut für Informatik, 2009.