C

Spreadsheet Modeling of Data Center Hotspots

E.T.T. Wong
Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hong Kong

M.C. Chan
Faculty of Business, The Hong Kong Polytechnic University, Hong Kong

L.K.W. Sze
Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong

INTRODUCTION

Today's data centers are evolving rapidly to embrace new technologies like the cloud, virtualization, consumer devices and more. As a result, data centers consume more energy than ever. A common management practice is that every data center periodically replaces its servers and major hardware during its technology refresh cycle. This permits organizations to apply updated technology to improve equipment reliability and enable new capabilities. Traditional data center refresh cycles took place about every five years, but owing to advanced technology, new guidelines for data processing environments, and alternative cooling technologies, they have been accelerating over the last decade.

In the case of advanced technology, a recent customer survey (Forrsights Hardware Survey, Q3 2011) revealed that among 161 enterprise IT infrastructure decision-makers across the US, the UK, and China, 57% have already adopted blade servers. Of those that have blade servers, nearly two-thirds are expanding their implementations (Forrester Consulting, 2012). Most users of blade servers reported increased server density and reduced power and cooling as the main operational and financial benefits. It can be anticipated that rapid adoption of refined blade technology will be forthcoming.

With regard to data processing environments, recently the American Society for Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) Technical Committee 9.9 described two new classes of data center equipment for which the maximum allowable temperature range is now 41 to 113°F (5 to 45ºC). This is a huge increase over ASHRAE's 2008 recommended range, which listed the maximum allowable temperatures at 59 to 89.6°F. The move to higher operating temperatures means that there is a significant reduction in cooling needs. Hence more and more operators would prefer to have the enhanced equipment in place as soon as possible.

However, many data centers are not expanding physically at the same rate as their computing power. As a result the load density per square foot is increasing in virtually all data centers and this will continue into the foreseeable future. Many data centers are now in the 20-40 watts per square foot with some already approaching 100 watts per square foot. The increased heat load and heat density takes place in the form of more powerful servers located in the same rack configuration, and in the same physical space as before. This is because data center operators and owners cannot afford to relocate on a regular basis as property prices are increasing. The emerging problem is how to manage the increased heat density within the constraints of the existing data center and a promising approach to ensure the achievement of various availability tier targets is to model data center hardware defects and failures.

For most data centers, huge centralized air conditioning units that push air through drop ceilings or raised floors remain a regular phenomenon, but for enterprises building out for energy efficiency or seeking to retrofit for added energy relief, localized cooling—mainly in the form of in-row cooling systems—is making a start.

In-row cooling systems are finding their place between racks, pumping out cold air through the front and pulling in hot air from the back. Because cooling is

DOI: 10.4018/978-1-4666-5888-2.ch115

performed by units just inches away from the heat source rather than indiscriminately through the floor or ceiling, data center hotspots run less hot. What is more, rather than relying on a central thermostat, these units function autonomously, tapping temperature-monitoring leads placed directly in front of a heat source to ensure that the air remains within a specified temperature range. Moreover, the unit ratchets down its cooling activities during partial loads. Nowadays, the cost-cutting benefits of localized cooling are quickly proving convincing, so much so that Gartner predicts in-rack and in-row cooling will become the predominant cooling method for the data center throughout the world.

However, as heat dissipation in data centers rises by orders of magnitude, inefficiencies such as air recirculation causing hotspots and flow short-circuiting has a significant impact on the system reliability and energy efficiency of the data center. Therefore, location and temperature distribution of hotspots is a significant challenge for data center management.

With the foreseeable mix of existing and new servers, there will be a change in the server room heat dissipation pattern associated with the original data center design and this change would likely lead to an entirely different distribution of hotspots. This article is aimed to propose a practical approach for modeling the temperature variations of hotspots in data centers. The following sections will provide key concepts of data center cooling, common causes of data center hotspots, and a description of the proposed modeling methodology. Finally the methodology will be illustrated with a real life application.

BACKGROUND

For high equipment reliability, identification of the hotspot locations and its temperature variations can significantly reduce the effect of unnecessary cooling energy. In general the cooling system for a data center comprises a Computer Room Air Conditioning (CRAC) unit and the associated air distribution system. A Computer Room Air Handling (CRAH) unit may be used instead of a CRAC in larger data centers. Nowadays, the appropriate cooling solution for a data center may be a combination of the following methods.

Methods of Cooling

- **Evaporative:** Evaporative cooling works by evaporating water into the air, very much like a human body coming out of a pool; the body always feels cooler because of the water being evaporated. An evaporative cooler must be used with adequate ventilation. The amount of cooling depends on the area's humidity. In humid environments cooling will average 8 to 12 degrees Fahrenheit and in dry environments 20 to 28 degrees Fahrenheit. This type of cooling can alleviate heat problems in large areas but is not a good choice for most server rooms but may serve a purpose in very large data centers.
- **Split cooling:** Split cooling consists of two components, a room unit and an outside condenser. Sometimes this is referred to as direct expansion cooling. The units are connected together by a refrigeration pipe that works with refrigerant. The external unit may take one of three forms:
  - A heat pump, which consists of a fan, compressor, coil, and reversing valve, rejects unwanted heat into the atmosphere during the cooling cycle but can also extract heat from the atmosphere during the heating cycle.
  - A condensing unit without a reversing valve provides cooling only.
  - A condenser, which consists of a fan and coil. The compressor is contained in the indoor unit. This method is used less often than the other split-cooling methods.
- **Water (or liquid) cooled.** These are refrigeration systems that chill liquids, such as water, oil, brine, alcohol, or various air conditioning chemicals. In air conditioning systems, the devices cool the liquid, typically water, and the cooled water chills and dehumidifies incoming circulated air. The system then releases the chilled air back into the atmosphere.

Related Content