Sustainable Computing: Informatics and Systems 19 (2018) 14–28

Contents lists available at ScienceDirect

Sustainable Computing: Informatics and Systems

journal homepage: www.elsevier.com/locate/suscom

A review of thermal management and innovative cooling strategies for data center

Chayan Nadjahi[a], Hasna Louahlia[a,*], Stéphane Lemasson[b]

[a] NORMANDIE UNIV, UNICAEN, LUSAC, Site universitaire Bellevue, 120 rue de l'Exode, 50000 Saint-Lô, France
[b] Orange Labs, 2, avenue Pierre Marzin, 22307 Lannion, France

ARTICLE INFO

Keywords:
Data center
Cooling technologies
Energy efficiency
Thermal management

ABSTRACT

Within a data center, 52% of the electricity is used by the IT equipment, 38% by the cooling system, and 10% for the remaining equipment (electrical power distribution, UPS…). The cooling system is mandatory to maintain optimal air temperature and relative humidity and provide a safe environment for all the electronic equipment. This paper presents a review on thermal management in data centers and various potential cooling technologies developed respecting energy saving constraint. Numerous researchers have proposed different alternative cooling strategy and solutions to free air cooling. This article describes a state of the art cooling strategies of data centers. The open aisle configuration is the most often used technique to remove heat. It does not affect energy efficiency but increases the electricity consumption dedicated to air conditioning. The most promising cooling systems, capable of absorbing energy flux while decreasing electricity consumption and rising the energy efficiency, are identified and discussed. It includes free cooling, liquid cooling, two-phase technologies and building envelope.

1. Introduction

Data center is a large-capacity facility (up to 500 000 m^2) in which are gathered Information Technology (IT) equipment, such as servers or processors, and support systems designed to provide a safe and reliable environment for IT equipment. An example of data center is shown in Fig. 1. Electricity used by data centers worldwide doubled from 2000 to 2005, then increased by about 56% from 2005 to 2010. Recent energy statistics indicate that the data center industry is responsible for 1.3% of the world and 2% of the United States electricity consumption [1]. In fact, between 25% and 35% of the worldwide power consumption of data centers is consumed by US data centers [1].

Within a data center, roughly 52% of the electricity is used by the information technology (IT) equipment, 38% by the cooling system, and 10% for the remaining equipment (electrical power distribution, UPS…). Most of the energy supplied to the IT equipment is converted into heat. Thus, the cooling system is necessary to maintain optimal air temperature and relative humidity and provide a safe environment for all the electronic equipment.

In a data center, the temperature maintenance is often handled by a Computer Room Air Conditioning (CRAC) unit. With this unit, it is important to configure air flows well in order to avoid that temperature rises too high, which could imply property damages. The most regularly used configuration is the open aisle configuration where cold air from air conditioning and hot air from racks are separated [2]. This configuration does not affect the energy efficiency but increases the electricity consumption dedicated to air conditioning [3]. That is why it is relevant to design a less energy-intensive cooling system.

This review paper introduces the four most promising cooling systems capable of decreasing the electricity consumption and rising the energy efficiency. The free cooling technology consists of using the natural fluid to cool data centers. The liquid cooled technology is useful when the data centers have a high-power density. The two-phase flow technology consists of using a refrigerant capable of removing dissipated heat by racks and rejecting to the outside environment. The building envelopes technology consists of using a container or phase change materials (PCMs).

2. Energy and environment context

2.1. Energy consumption

The repartition of the electricity consumption in a data center is specified by Fig. 2. Demand-side systems, including processors, server power supplies, or storage and communication equipment, account for 52% of total consumption. Supply-side systems include the UPS, power

* Corresponding author.
E-mail address: hasna.louahlia@unicaen.fr (H. Louahlia).

https://doi.org/10.1016/j.suscom.2018.05.002
Received 24 October 2017; Received in revised form 14 March 2018; Accepted 15 May 2018
Available online 24 May 2018

Fig. 1. Data center in Nokia, Lannion.

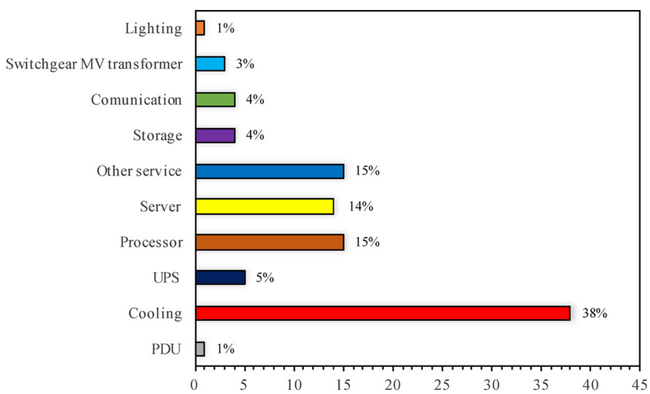

Fig. 2. Electricity distribution in data centers [4].

distribution, lighting and building switchgear represent 10% of total consumption. Energy consumption of data center traditional cooling is around 38% of data center total energy consumption [4].

The worldwide electricity usage in data centers has increased between 2000 and 2005 from 71 to 152 billion kWh per year, as shown in Fig. 3. This growth of the electricity consumption represents approximately 10% per year. In 2005, electricity consumed by data centers was about 1% of world electricity use. By 2010, world data center electricity use represents between 1.1 and 1.5% of the world electricity use. The lower bound figures represent 20 to 33% growth in data center electricity use compared to 2005. According to the Japanese Ministry of Economy, the electricity consumption will be five times greater in 2025 [6]. The strong electricity usage, particularly in cooling, has placed energy efficiency at the top of the agenda for both datacom businesses and policy makers [5].

2.2. Carbon footprint

The environmental impact has to take into account: in 2008, the Smart 2020 study highlights that in 2002, the global data center footprint, including equipment use and embodied carbon, was 76 $MtCO_2e$ and, by 2020, the value of 259 $MtCO_2e$ is expected – an increase of 7% per year until 2020 [7]. Fig. 4 shows the composition of data center footprint in 2002 and 2020.

2.3. Metrics for energy efficiency of data centers

The Power Usage Effectiveness (PUE) is proposed by the Green Grid initiative [8] as a fraction of total power of the data center to that used

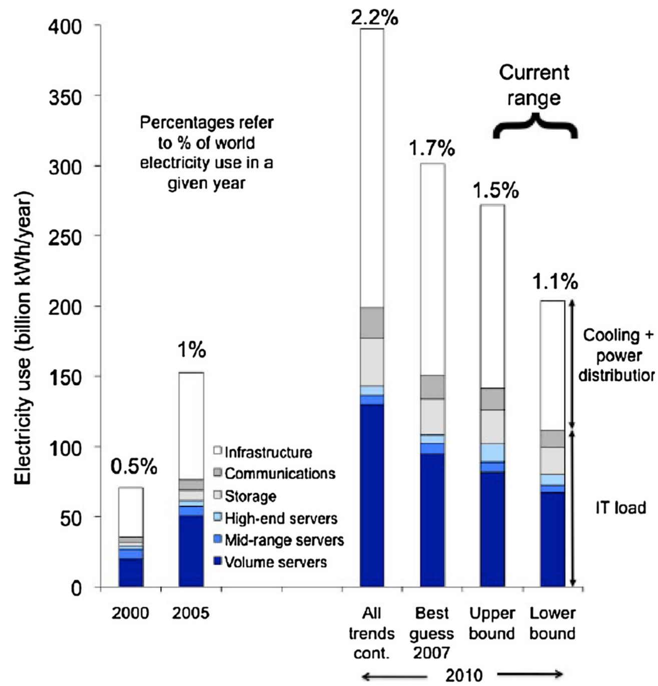

Fig. 3. Electricity consumption worldwide in data centers [6].

by the IT equipment. It is defined to assess the energy efficiency of data center over the year:

$$PUE = \frac{P_{DC}}{P_{IT}} \tag{1}$$

Where P_{DC} is the power of the data center, P_{IT} is the input power of the IT equipment

This can be expanded as [9]:

$$PUE = \frac{P_{cooling} + P_{power} + P_{lighting} + P_{IT}}{P_{IT}} \tag{2}$$

Where $P_{cooling}$ is the cooling equipment input power, P_{power} is the power lost in the energy distribution system through line-loss and other infrastructure (UPS or PDU) inefficiencies, $P_{lighting}$ is the power used to light the data center and support spaces, P_{IT} is the input power of the IT equipment.

A PUE of 1 would be an ideal value because it implies that all energy is used by the IT equipment. However this is possible if there is no cooling equipment or no power delivery components [10]. With proper design, a PUE value of 1.6 should be achievable but it could be equal to 1.2 [4]. The reduction in this value depends strongly on the cooling design and its effectiveness. PUE could be reduced by 50% using liquid cooling (passive or active) instead of a traditional raised floor. The reduction of PUE leads to increase the power per cabinet from 5 kW to 40 kW [5]. A study conducted by Lawrence Berkeley National Labs [11] where the authors benchmarking 22 data centers and showed a set of the best practice technologies for reducing the PUE value among which the authors included the use of evaporative liquid cooling and energy optimization of the cooling infrastructure.

The PUE itself does not take into account any reuse of energy, therefore the Green Grid defined the Energy Reuse Effectiveness (ERE) as [9]:

$$ERE = \frac{P_{cooling} + P_{power} + P_{lighting} + P_{IT} - P_{reuse}}{P_{IT}} \tag{3}$$

Where $P_{cooling}$ is the input power of the cooling equipment, P_{power} is the power energy lost in the power distribution system through line-loss and other infrastructure (UPS or PDU) inefficiencies, $P_{lighting}$ is the power used to light the data center and support spaces, P_{IT} is the input