On the Leakage-Power Modeling for Optimal Server Operation

Patricia Arroba [1], Marina Zapater [2], José L. Ayala[3], José M. Moya[1], Katzalin Olcoz[3] and Román Hermida[3]

Resumen— **Leakage power consumption is a component of the total power consumption in data centers that is not traditionally considered in the setpoint temperature of the room. However, the effect of this power component, increased with temperature, can determine the savings associated with the careful management of the cooling system, as well as the reliability of the system. The work presented in this paper detects the need of addressing leakage power in order to achieve substantial savings in the energy consumption of servers. In particular, our work shows that, by a careful detection and management of two working regions (low and high impact of thermal-dependent leakage), energy consumption of the data-center can be optimized by a reduction of the cooling budget.**

Palabras clave— **Power consumption, Leakage, cooling, efficiency.**

I. Introduction

ONE of the big challenges in data centers is to manage system resources in a power-efficient way. Data centers consume from 10 to 100 times more power per square meter than typical office buildings [1]. They can even consume as much electricity as a city [2]. The power consumption budget in data centers comes from computation processing, disk storage, network, and cooling systems.

It must be said that greening the computer industry is touching off an unprecedented level of cooperation and information-sharing among companies, government, and laboratories. In the USA, *Green Grid*, a consortium of industry leaders (like AMD, Intel, Dell, HP, IBM, Sun, and others who are normally competitors) to share data and strategies for greener data centers. Green Grid's membership also includes the Pacific Gas and Electric Company (better known as PGE), and it recently announced a collaboration agreement with the U.S. Department of Energy.

However, data center designers have collided with the lack of accurate power models for the energy-efficient provisioning of their devised infrastructures, and the real-time management of the computing facilities. The work proposed in this paper makes substantial contributions in the area of power modeling of high-performance servers for data center-operated services.

Interestingly, the key issue of how to control the setpoint temperature at which to run the cooling system of a data center, is still to be clearly defined [3]. Data centers typically operate in a temperature range between 20° C and 22° C, but we can find some of them as cold as 13° C degrees [4,5]. Due to lack of scientific data in the literature, these values are often chosen based on conservative suggestions provided by the manufacturers of the equipment. Some authors estimate that increasing the setpoint temperature by just one degree can reduce energy consumption by 2 to 5 percent [4,6]. Microsoft reports that raising the temperature from two to four degrees in one of its Silicon Valley data centers saved $250,000 in annual energy costs [5]. Google and Facebook have also been considering increasing the temperature in their data centers [5].

Power consumption in servers can be estimated by the summation of the dynamic power consumption of every active module, dependent on the activity, and the leakage power consumption, that is strongly correlated with the integration technology. In particular, leakage power consumption is a component of the total power consumption in data centers that is not traditionally considered in the setpoint temperature of the room. However, the effect of this power component, increased with temperature, can determine the savings associated with the careful management of the cooling system, as well as the reliability of the system itself.

The work presented in this paper detects the need of addressing leakage power in order to achieve substantial savings in the energy consumption of servers and makes the following contributions:

- we establish the need of considering leakage power consumption and its dependency with temperature for modern data centers;
- we detect and define two working regions depending on the impact of leakage power in the total power consumption of high-performance servers;
- we observe that substantial energy savings can be achieved by the careful management of the cooling system once the previous contribution has been verified;
- we validate the previous hypothesis with a deep experimental work that resembles the infrastructure of current enterprises.

II. Related work

In [7] a statistical model that provides run-time system-wide prediction of energy consumption on server blades is proposed. The authors develop a

[1]Electronic Engineering Dept., ETSI. Telecomunicación, Universidad Politécnica de Madrid, e-mail: {parroba,josem}@die.upm.es
[2]CEI Campus Moncloa UCM-UPM, e-mail: marina@die.upm.es
[3]DACYA, Universidad Complutense de Madrid, e-mail: {jlayalar,katzalin,rhermida}@ucm.es

linear regression model that relates processor power, bus activity, and system ambient temperatures into real-time predictions of the power consumption. Other works such [8–10] also present the power consumption of a server as a linear function of the CPU usage of that server.

Some other linear models can be found in [9], where server's power is formulated as a quadratic function of the CPU usage, or in [11], where the transition between server functionality state (Idle - ON) is taken into consideration. The work in [12] follows a similar approach but, in this case, the CPU power consumption percentage is separated in two parts: the one due to the applications, and the second one due to management services turning on and off the server. The power modeling technique vMeter, proposed by Bohra et al. [13] observes a correlation between the total system's power consumption and component utilization. They created a four-dimensional linear weighted power model for the total power consumed $P(total)$ by separating the contribution of each active domain in a node. However, none of these works have considered the effect of leakage power and temperature in the total power consumption of the servers.

The work presented in [14] is most relevant for us. In this paper, the authors compare the impact of increasing the air temperature entering the rack on the complete cooling infrastructure, with the alternative approach of allowing greater air temperature rise across the rack. Even though in their approach they develop a power model from chip to cooling tower, they still ignore the leakage as a key factor in data center power consumption and energy saving opportunities.

III. Background on data center power modeling

The main contributors to the energy consumption in a data center are the computing power (also known as IT power), i.e. the power drawn by servers in order to run a certain workload, and the cooling power needed to keep the servers within a certain temperature range that ensures safe operation. Traditional approaches have tried to reduce the cooling power of data center infrastructures by increasing the supply temperature of Computer Room Air Conditioning Units (CRAC units). However, because of the direct dependency of leakage current with temperature, the leakage-temperature tradeoffs at the server level must be taken into account when optimizing energy consumption.

In this section we show the impact of these tradeoffs on the total energy consumption of the data center, as well as how the ambient room temperature influences the cooling power of data centers. This fact, as will be shown later, can be exploited to optimize the power consumption of the data center.

A. Computing power

Current state-of-the-art resource management and selection techniques were contemplating only the dynamic power consumption of servers when allocating tasks or selecting machines. Moreover, the devised power models have not traditionally included the impact of leakage power consumption and its thermal dependency, driving to non-optimal solutions in their energy optimization plans.

Theoretically, no electrical current should circulate through the substrate of a MOS transistor between drain and source when it is powered off due to an infinite gate resistance. However, in practice this is not true, and leakage currents flow through the reverse-biased source and drain-bulk pn junctions in dynamic logic. Also due to the continuous technology scaling, the influence of leakage effects is rising, increasing junction leakage currents by 5 orders of magnitude compared to previous feature sizes according to Rabaey [15].

Dynamic consumption has historically dominated the power budget. But when scaling technology below the $100nm$ boundary, static consumption becomes much more significant, being around 30-50% [16] of the total power under nominal conditions. This issue is intensified by the influence of temperature on the leakage current behavior. With increasing temperature the on-current of a transistor is reduced slightly. However, the reduction of the threshold voltage is not sufficient to compensate for the decreased carrier mobility that has a strong exponential impact on leakage current.

Therefore, it is important to consider the strong impact of static power consumed by devices as well as its temperature dependence and the additional effects influencing their performance. In this section, we derive a leakage model for the static energy consumption of servers and we validate it with real measurements taken in an AMD Opteron machine of our case study.

The current that is generated in a MOS device due to leakage is the one shown in equation 1.

$$I_{leak} = I_s \cdot e^{\frac{V_{GS}-V_{TH}}{nkT/q}} \cdot (1 - e^{\frac{Vds}{kT/q}}) \tag{1}$$

Research by Rabaey [15] shows that if $V_{DS} > 100mV$ the contribution of the second exponential is negligible, so the previous formula can be rewritten as in Equation 2:

$$I_{leak} = I_s \cdot e^{\frac{V_{GS}-V_{TH}}{nkT/q}} \tag{2}$$

where technology-dependent parameters can be grouped together to obtain the formula in Equation 3:

$$I_{leak} = B \cdot T^2 \cdot e^{\frac{V_{GS}-V_{TH}}{nkT/q}} \tag{3}$$

where B defines a constant that depends on the manufacturing parameters of the server.

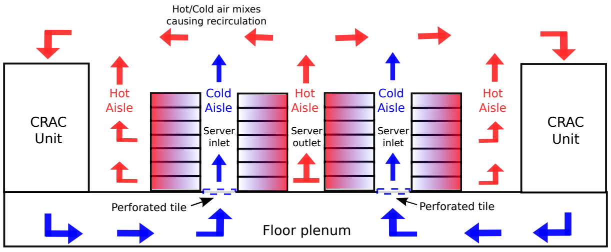

Fig. 1. Data Center cooling scheme

B. *Cooling power*

The cooling power is one of the major contributors to the overall data center power budget, consuming over 30% of the overall electricity bill in typical data centers [14]. In a typical air-cooled data center room, servers are mounted in racks, arranged in alternating cold/hot aisles, with the server inlets facing cold air and the outlets creating hot aisles. The CRAC units pump cold air into the data room and extract the generated heat (see Figure 1). The efficiency of this cycle is generally measured by the *Coefficient of Performance* (COP). The COP is a dimensionless value defined as the ratio between the cooling energy produced by the air-conditioning units (i.e. the amount of heat removed) and the energy consumed by the cooling units (i.e. the amount of work to remove that heat), as shown in Equation 4.

$$COP_{MAX} = \frac{\text{output cooling energy}}{\text{input electrical energy}} \tag{4}$$

Higher values of the COP indicate a higher efficiency. The maximum theoretical COP for an air conditioning system is described by Carnot's theorem as in Equation 5:

$$COP_{MAX} = \frac{T_C}{T_H - T_C} \tag{5}$$

where T_C is the cold temperature, i.e. the temperature of the indoor space to be cooled and T_H is the hot temperature, i.e. the outdoor temperature (both temperatures in Celsius). As the difference between hot and cold air increases, the COP decreases, meaning that the air-conditioning is more efficient (consumes less power) when the temperature difference between the room and the outside is smaller.

According to this, one of the techniques to reduce the cooling power is to increase the COP by increasing the data room temperature. We will follow this approach to decrease the power wasted on the cooling system to a minimum, while still satisfying the safety requirements of the data center operation.

IV. Experimental methodology

The experimental methodology in this paper pursues two goals: (i) to describe the leakage-temperature tradeoffs at the server level, by means of measuring the power consumption of an enterprise server at different temperatures and under a controllable workload; and (ii) to validate the model in a real data room environment where the air-conditioning can be controlled. After this, we will be able to evaluate the energy savings that could be obtained in a data center when our modeling strategy is applied.

A. *Server-level setup*

As temperature-dependent leakage cannot be measured separately from the dynamic power in a server by the power measurement devices, we use a controllable workload, *lookbusy*[1], in order to explore the leakage at the server level. *Lookbusy* can stress all the hardware threads to a fixed CPU utilization percentage without memory or disk usage, for a paricular period of time. The usage of a synthetic workload to derive the leakage model has many advantages, the most important of which is that dynamic power can be described as linearly dependent with CPU utilization and Instructions Per Cycle (IPC), or kept constant. In our case, we stress the system at the maximum CPU utilization, in order to isolate the dynamic power. The platform under test is a SunFire V20z server with 2 Dual-Core AMD Opteron processors and 4GB of RAM. Keeping the workload constant, we slowly vary the inlet temperature of the server, obtaining CPU temperatures ranging from 45°C to 70°C, while monitoring the following server parameters: (i) CPU temperature (1 sensor per CPU), (i) memory temperature (1 sensor per each of the 2 memory banks), (iii) fan speed and (iv) overall power consumption.

All temperatures and fan speed values are obtained via the server internal sensors, collected through the Intelligent Platform Management Interface (IPMI) tool [2]. IPMI allows to poll the internal sensors of the enterprise server with negligible overhead. Because the server is not shipped with power consumption sensors, we use non-intrusive current clamps. The current clamp is connected to the power cord of the server and wirelessly transmits the monitored data to a base station connected to a desktop computer. Data gathered via the current clamp and the server internal sensors are aligned to ensure they have a common timestamp.

Our hypothesis is that we can find and define two different working regions depending on the impact of leakage power in the total power consumption of high-performance servers. In this sense, we aim to prove that for the lower range of temperatures, the impact of the temperature-dependant leakage is negligible, whereas for a higher temperature range leakage needs to be considered. This hypothesis will be verified throughout the extensive experimental work and the methodology just described.

B. *Data room setup*

In order to validate the server-level model in an infrastructure resembling a real data center scenario, we install eight Sunfire V20z servers in a rack inside

[1] http://www.devin.com/lookbusy/
[2] http://ipmitool.sourceforge.net/

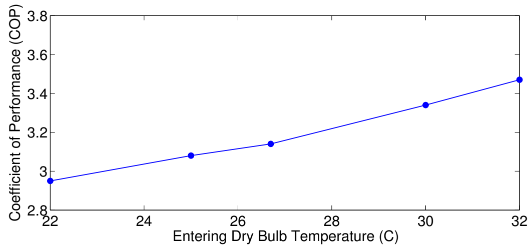

Fig. 2. Evolution of the air-conditioning COP with room temperature

an air-cooled data room, with the rack inlet facing the cold air supply and the outlet facing the heat exhaust. The air conditioning unit mounted in the data room is a Daikin FTXS30 unit, with a nominal cooling capacity of 8.8kW and a nominal power consumption of 2.8KW. We assume an outdoor temperature of 35°C and use the manufacturers technical data to obtain the COP curve depending on the room temperature [17]. This temperature is only used to estimate the energy savings based on the curve provided by the manufacturer and does not affect the experimental results.

As can be seen in Figure 2, as the room temperature and the heat exhaust temperature raises, approaching the outdoor temperature, the COP increases and, thus the cooling efficiency improves.

We monitor all the servers by means of IPMI tool to gather the server internal sensors and the current clamps to obtain power consumption. We set the air supply temperature at various values ranging from 18°C to 24°C, and run from 1 to 4 simultaneous instances of the different tasks of the SPEC CPU 2006 benchmark suite [18] in the servers of the data room. Our goal is to verify the leakage-temperature model, finding the maximum air-supply temperature that makes the servers work in the temperature region where leakage is negligible.

V. Results

Characterizing the power with respect to temperature under a constant synthetic workload allows us to define different working regions depending on the impact of leakage power. In region I, for CPU temperatures ranging from 44° C to 48° C, we find that the contribution of the leakage to the total consumption of the server is negligible (see Figure 3). As can be seen, the obtained data follows a linear trend, as expected. Once the regression model is built, we obtain Equation 6 that fits the experimental data.

$$P_I = 0.0288 \cdot T_{CPUaverage} + 182.15 \qquad (6)$$

On the other hand, region II is defined for those CPU temperatures higher than 48° C. In this region, the impact of power consumption due to leakage needs to be considered, as can be seen in Figure 4. The data fitting to a linear curve in this region is shown in Equation 7, where an increase of about one order of magnitude in the slope of the curve can be appreciated if compared with region I.

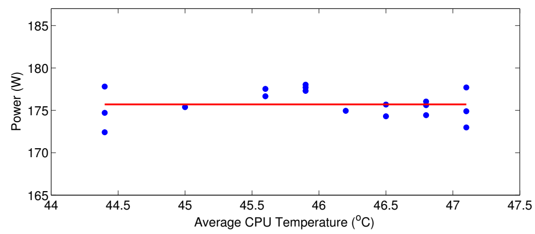

Fig. 3. Power consumption of SunFire V20z for temperature region I

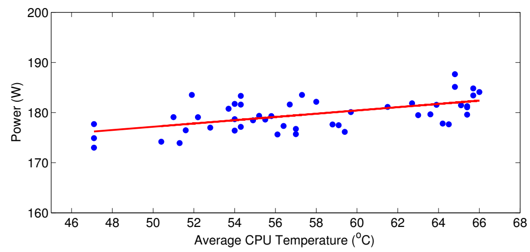

Fig. 4. Power consumption of SunFire V20z for temperature region II

$$P_{II} = 0.3255 \cdot T_{CPUaverage} + 160.894 \qquad (7)$$

In both cases, the dispersion of the samples are due to the inaccuracy of the clamp, whose error in performing the measurements is close to ±5W.

After obtaining the two working regions for the leakage power, we move to the data room setup. We run the tasks of the SPEC CPU 2006 benchmark suite in the AMD servers under different data room conditions. In our experiments, we run from 1 to 4 instances of SPEC CPU in the AMD servers at different room temperatures of 18° C, 20° C, 22° C and 24° C. Figure 5a shows the power consumption values for two simultaneous instances of the SPEC CPU 2006 benchmark at an air supply setpoint temperature of 18° C, 20° C and 24° C, respectively. Figure 5b shows the CPU temperature for each of these tests under the same conditions.

Because all other variables are constant, and as

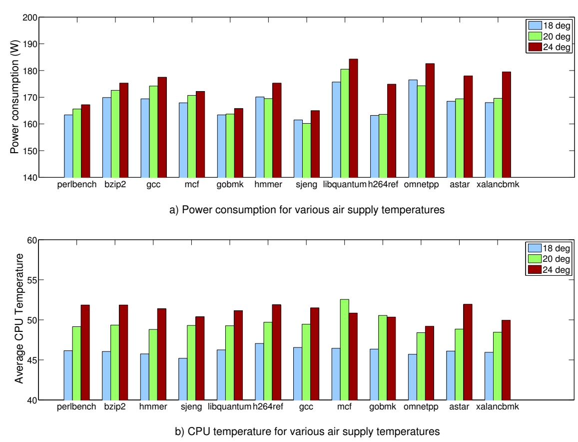

Fig. 5. Power consumption of SPEC CPU 2006 at different air supply temperatures

the measurement error with the current clamp is already controlled, the changes in the power consumption for each test can be due to the differences in ambient temperature. As can be seen in the plots, even though there are differences in the average CPU temperature between the 18° C and the 20° C case, for most of the benchmarks CPU temperature does not go above the 50° C, staying in the negligible leakage area. In fact, the power consumption differences between the 18° C and the 20° C case are in the range of ±5W, so we cannot consider them to be due to leakage, but to the inaccuracy of our current clamp. However, for the 24° C case, CPU temperatures raise above 50° C and power consumption for most of the benchmarks is considerably higher than in the 18° C scenario, achiving differences higher than 8W for gcc, libquantum, astar and xalancbmk benchmarks. Thus, in this region we begin to observe temperature-dependant leakage.

The experimental results for our data room scenario show that if we allow temperature to raise above this 24° C barrier, the contribution of the leakage increases, increasing the computing power drawn by our infrastructure. However, for our data room configuration and under our workload, leakage is negligible in the 18° C-24° C range and, thus, we can raise the ambient temperature in order to reduce cooling power.

If we increase the air supply temperature from 18° C to 24° C, the room temperature increases and the COP varies (see Figure 2) from 2.95 to 3.47, increasing the energy efficiency of the cooling equipment and reducing the cooling power. This increase has a proportional impact on the energy savings of the infrastructure, leading to a decrease of 11.7% in cooling power as predicted by the curve.

VI. Conclusions

Power consumption in servers can be estimated by the summation of the dynamic power consumption of every active module, dependent on the activity, and the leakage power consumption, that is strongly correlated with the integration technology. However, traditional approaches have never incorporated the impact of leakage power consumption in these models, and the noticeable values of leakage power consumption that appear at higher CPU temperatures.

The work presented in this paper detects the need of addressing leakage power in order to achieve substantial savings in the energy consumption of servers. In particular, our work shows that, by a careful detection and management of two working regions (low and high impact of thermal-dependent leakage), energy consumption of the data-center can be optimized by a reduction of the cooling budget. Finally, we validate these facts with a deep experimental work that resembles the infrastructure of current enterprises, where an 11 % of the cooling budget can be reduced.

Acknowledgement

Research by Marina Zapater has been partly supported by a PICATA predoctoral fellowship of the Moncloa Campus of International Excellence (UCM-UPM). This work has been partially supported by the Spanish Ministry of Economy and Competitiveness, under contracts TIN2008-00508, TEC2012-33892 and IPT-2012-1041-430000, and INCOTEC. The authors thankfully acknowledge the computer resources, technical expertise and assistance provided by the Centro de Supercomputación y Visualización de Madrid (CeSViMa).

Referencias

[1] P. Scheihing, "Creating energy efficient data center," in *Data Center Facilities and Engineering Conference*, Washington DC, USA, May 2007.

[2] J. Markoff and S. Lohr, "Intel's huge bet turns iffy," *New York Times Technology Section*, September 2002.

[3] Nosayba El-Sayed, Ioan A. Stefanovici, George Amvrosiadis, Andy A. Hwang, and Bianca Schroeder, "Temperature management in data centers: why some (might) like it hot," in *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems*, New York, NY, USA, 2012, SIGMETRICS '12, pp. 163–174, ACM.

[4] J. Brandon, "Going green in the data center: Practical steps for your SME to become more environmentally friendly," *Processor*, , no. 29, 2007.

[5] Rich Miller, "Google: Raise your data center temperature.," October 2008.

[6] "Summer time energy-saving tips," .

[7] Adam Lewis and et al., "Run-time energy consumption estimation based on workload in server systems," in *HotPower*, Berkeley, CA, USA, 2008, pp. 4–4.

[8] Steven Pelley and et al., "Understanding and abstracting total data center power," in *WEED*, June 2009.

[9] Xiaobo Fan and et al., "Power provisioning for a warehouse-sized computer," in *ISCA*, New York, NY, USA, 2007, pp. 13–23.

[10] Frank Bellosa, "The benefits of event: driven energy accounting in power-sensitive systems," in *ACM SIGOPS*, New York, NY, USA, 2000, pp. 37–42.

[11] David Meisner and et al., "Peak power modeling for data center servers with switched-mode power supplies," in *ISLPED*, New York, NY, USA, 2010, pp. 319–324.

[12] G. Warkozek and et al., "A new approach to model energy consumption of servers in data centers," in *ICIT*, 2012, pp. 211–216.

[13] A.E.H. Bohra and V. Chaudhary, "Vmeter: Power modelling for virtualized clouds," in *IPDPSW*, 2010, pp. 1–8.

[14] T.J. Breen, E.J. Walsh, J. Punch, A.J. Shah, and C.E. Bash, "From chip to cooling tower data center modeling: Part i influence of server inlet temperature and temperature rise across cabinet," in *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2010 12th IEEE Intersociety Conference on*, 2010, pp. 1–10.

[15] J. Rabaey, *Low Power Design Essentials*, Engineering (Springer-11647). Springer, 2009.

[16] S.G. Narendra and A.P. Chandrakasan, *Leakage in Nanometer CMOS Technologies*, Integrated Circuits and Systems. Springer, 2010.

[17] Daikin AC (Americas), Inc., "Engineering data split, ftxs-l series," 2010.

[18] SPEC CPU Subcommittee and John L. Henning, "SPEC CPU 2006 benchmark descriptions," http://www.spec.org/cpu2006/.