12th IMEKO TC10 Workshop on Technical Diagnostics
New Perspectives in Measurements, Tools and Techniques for Industrial Applications
June 6-7, 2013, Florence, Italy

Evaluation of Power Consumption of Workstation Computers using Benchmarking

Francesco Picariello, Sergio Rapuano, Umberto Villano

University of Sannio, Corso Garibaldi, 107, Palazzo dell'Aquila Bosco Lucarelli, 82100 Benevento, Italy
{fpicariello, rapuano, villano}@unisannio.it

Abstract **-** This paper deals with the design and realization of a measurement system for estimating power consumption of workstation computers during the execution of benchmarks. The measurement system and the test bench architecture, designed to measure the instantaneous power consumption from the power line of a workstation motherboard, are presented. Then, the calibration procedure used to compensate systematic effects and to characterize the power measurement system is described. Five benchmarking applications have been performed on an Intel Pentium 4 processor and the corresponding results are shown and discussed in the paper.

I. Introduction

Nowadays, processors are designed for optimizing both performance and energy efficiency [1], especially for mobile applications where one of the most important issues is battery life. Over the years, the interest in the analysis of the power consumption on different processor platforms increased. Several papers have been published focusing on the development of analytical models to find the trade off between power and performance for different processor architectures. In [2], an analysis of power and performance for pipelined processors is presented. The authors propose a power consumption model for a generic out-of-order 8-way superscalar processor, called Turandot [3]. This model depends on the Fan-Out-of-four (FO4) delay of the pipeline length. In order to choose the optimal pipeline depth as a function of power and performance, the power consumption for different workloads is presented in the paper. The proposed model can be used to design low-power processors, but it provides the power consumption on basis of an ideal processor architecture model. The described paper highlights the needs of power consumption measurements in order to correlate them to the processor activities during the execution of workloads on several processor platforms. To achieve this aim, the instantaneous power consumption for each processor activity (floating-point operations, the writing and reading of the cache memory and the use of the ALU) should be measured. The benchmark applications, usually do not provide the estimation of the instantaneous power consumption of general-purpose processors at the micro-operation level because (i) this requires the knowledge of details of the hardware architecture that are kept secret by the manufacturers, (ii) no measurement of the actual power consumed is provided by the manufacturers at the operating system level, and (iii) the manufactures provide only the Thermal Design Power (TDP) and the Average CPU Power (ACP), which is the average power consumption of the processor while running a collection of 5 different benchmarks [4]. Therefore, it is necessary to embed a power measurement system on the power supply pins of the processor. Different measurements of performance and power consumption on low-power processors are presented in [5] for the following benchmark suites: Phoenix MapReduce, MiBench and SPEC CPU2000. By using a data-logger coupled with a current measuring probe, the performance and the energy-efficiency of three low power processors, Intel Atom D510, Freescale QorIQ P2020, and Texas Instruments DM3730 have been measured. In [6], the average chip power consumption and performance of five processors executing 61 benchmarks is presented. For each of them, the average power consumption was evaluated in four processor configurations (single and dual cores, single and dual threads). The power consumption of the processor was measured by using a current sensor based on the Hall effect, inserted on the processor power supply wire (12 V power line) from the computer motherboard. This power measurement system has the following limits: (i) does not measure the voltage on the 12 V line, presumed constant and equal to the nominal value; (ii) does not provide measurements that can be traced back to the International System of Quantities; and (iii) provides measurements with a fixed sampling frequency of 50 Hz that can be used only for measuring average power consumptions due to long list of instructions.

The long-term objective of the research described in this paper is to correlate the power consumption to the microscopic internal factors related to the execution of a given code (cache misses, mispredicted branches, number of active threads, page faults, use of floating point unit and of SIMD instructions, just to mention a few) with adjustable time resolutions going below 1 ms. The correlation between the power consumption and the processor activities can be used for: (i) the hardware design of low power processors, and (ii) the development of software applications oriented to the optimization of the power consumption. The first aim requires the development of power consumption models, which take into account the architectural details of the processor (e.g. depth of the pipeline, size of the cache memories and number of ALUs). The second aim requires the development of software tools oriented to evaluate the power consumption, on different processor platforms, for applications written in different programming languages.

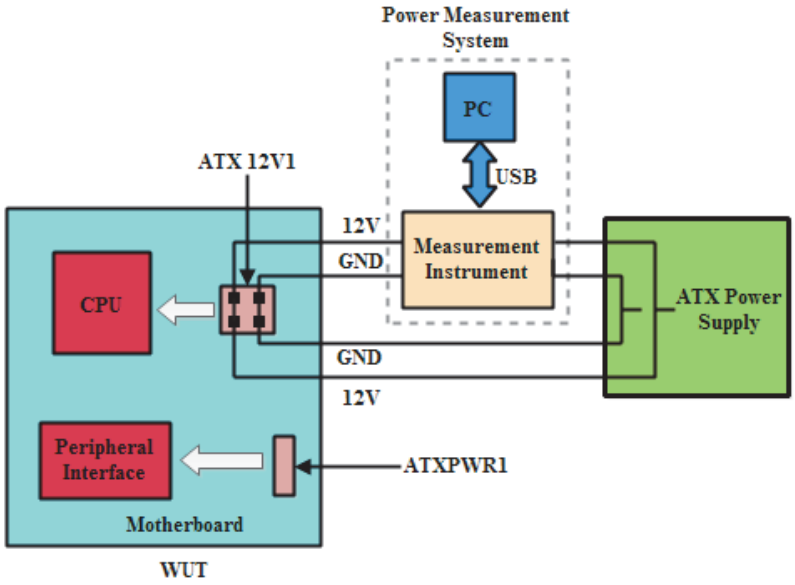

Figure 1. Test bench architecture.

To this aims, the steps to be performed are: (i) development of a mathematical model for evaluating the processor power consumption during the execution of different workloads, (ii) development of a software tools, which estimate the power consumption by using the developed mathematical model, and (iii) validation of the software tools for estimating the power consumption by comparing the results achieved by using contemporarily a power measurement instrument and the software tools.
As a first step, a power measurement system for measuring the instantaneous power consumption of a workstation computer is presented in this paper. The system has been calibrated to provide traceable measurements and tested on a single CPU type (even if a very popular one), running freely available benchmark codes. The rest of the paper is organized as it follows. The next section presents a general overview of the test bench architecture, the Workstation Under Test (WUT) and the power measurement system. Section III describes the benchmarks, which have been used during the measurements of the power consumption. The preliminary results and their discussion are presented in Section IV. The last section contains authors' conclusions.

II. Test bench architecture

In order to measure the power consumption of a WUT, it was necessary to design an embedded measurement instrument and a test bench. The proposed test bench is presented in Figure 1 and it is composed of: (i) the WUT motherboard, (ii) the power measurement system which measures the voltage and the current from the 12 V power line (ATX12V1) of the ATX connector, and (iii) the WUT ATX Power Supply. From Figure 1, it can be seen that the ATX12V1 connector presents two 12 V lines connected on parallel for supplying the motherboard. The measurement instrument is connected in series with one of the two lines coming from the ATX Power Supply connector. The power measurement system is composed of: (i) a current sensor and a voltage sensor, and (ii) a host PC, which receives the acquired samples by means of a Universal Serial Bus (USB) connection, calculates the instantaneous power consumption, displays and stores the results. In the following, the characteristics of the WUT, the architectural overview of the power measurement system and its calibration are described.

A. Workstation under test

The WUT considered for the power consumption measurements embeds an Intel Pentium 4 processor and a motherboard P4C800-E [7]. The motherboard is powered by the ATX 12 V power supply, by two connectors: the 20-pin ATXPWR1 for the peripheral interfaces and the 4-pin ATX12V connector (2 for ground and 2 for 12 V), which provides the power supply to the CPU (see Figure 1). As quoted above, separated lines supply the current on the 12V line. Thus, the total current consumption can be evaluated by multiplying the measured value of the current consumption on one wire by 2. The CPU under test is the Intel Pentium 4 Northwood, which is a single core processor with hyperthreading, running at 2798.72 MHz. The WUT has a Double Data Rate (DDR) memory of 512 MB with a memory clock of 199.9 MHz.

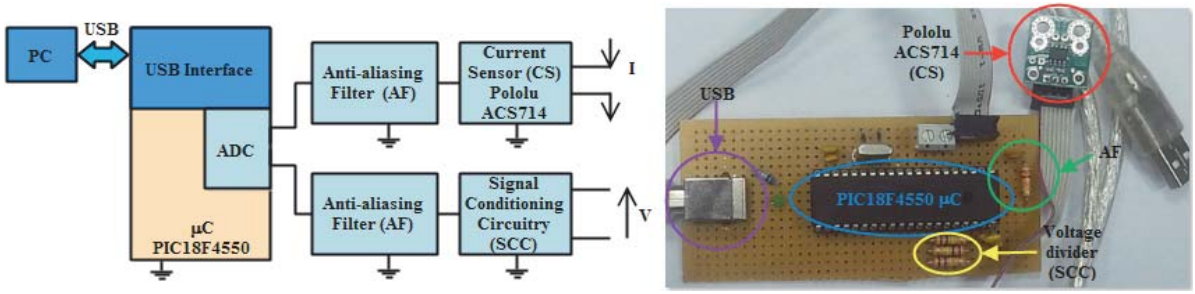

Figure 2. Measurement system architecture.

B. Power measurement system

The architecture of the designed power measurement system is presented in Figure 2. It consists of a PIC18F4550 microcontroller (µC) [8], used for data acquisition with a sampling frequency up to 50 kHz, two anti-aliasing filters, one current measurement sensor mounted on an evaluation board, Pololu ACS714, which exhibits a low input impedance of 1.2mΩ [9], a signal conditioning circuit and the host PC. The outputs from the current sensor (I) and the signal conditioning circuitry (V) were acquired by means of two analog inputs of the PIC18F4550 microcontroller. The µC includes a 10-bit Successive Approximation Register (SAR) Analog to Digital Converter (ADC). The Pololu ACS714 (CS) board measures the currents within a magnitude up to 5A and it is powered at 5V/10mA [9]. It provides an output voltage, which is proportional to the current on the input terminals, having a nominal output sensitivity of 185 mV/A. The signal conditioning circuitry (SCC) is composed of a voltage divider, which adapts the 12 V to the input measurement range of the microcontroller (5V). The voltage divider was implemented by using three resistors of 266 kΩ, in order to exhibit high input impedance (around 800 kΩ). Each of the anti-aliasing filters (AFs) is composed of a filter bank, which was designed to work at three different cut-off frequencies. The filters are placed between the input ADC pins of the microcontroller and the output pins of the current measurement sensor and the signal conditioning circuitry, respectively. The host PC manages the microcontroller by a set of commands, in order to start the data acquisition, to stop the data acquisition, and to set the sampling frequency (e.g., 50 Hz, 1 kHz or 50 kHz). The cut-off frequency of each filter is changed accordingly in the PIC. For this reason, a LabVIEW application has been developed in order to interact with the measurement system. It allows sending commands to the microcontroller, displaying the measured values to the user, storing them on the host PC and evaluating the power consumption from the voltage and the current values.

Figure 3 shows the flowchart of the power measurement software running on the µC. The µC software steps though the following phases: (i) initialization of the USB interface, (ii) reading from the USB buffer, (iii) waiting for the start of data acquisition command, (iv) reading the sampling frequency from the USB buffer, (v) initialization of the ADC, in order to acquire data from the ADC0 and ADC1 channels, at the selected sampling frequencies, (vi) if the µC receives the command of stopping data acquisition, it returns to the USB initialization, otherwise it continues the data acquisition, (vii) polling the end of conversion flag, which is high when the ADC has converted the signal from the ADC0 channel, (viii) reset of the flag, (ix) reading of the ADC0 channel value, (x) polling the flag in case of ADC1 channel, (xi) reset of the flag, (xii) reading of the ADC1 channel value, and (xiii) sending the acquired data via USB to the host PC.

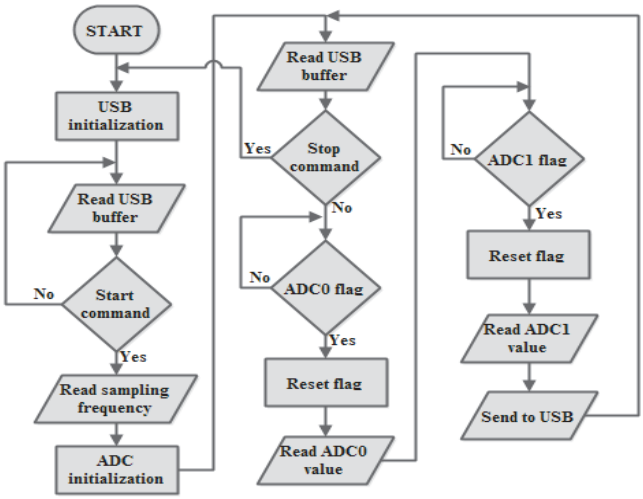

Figure 3. Flowchart of the software running on the microcontroller.

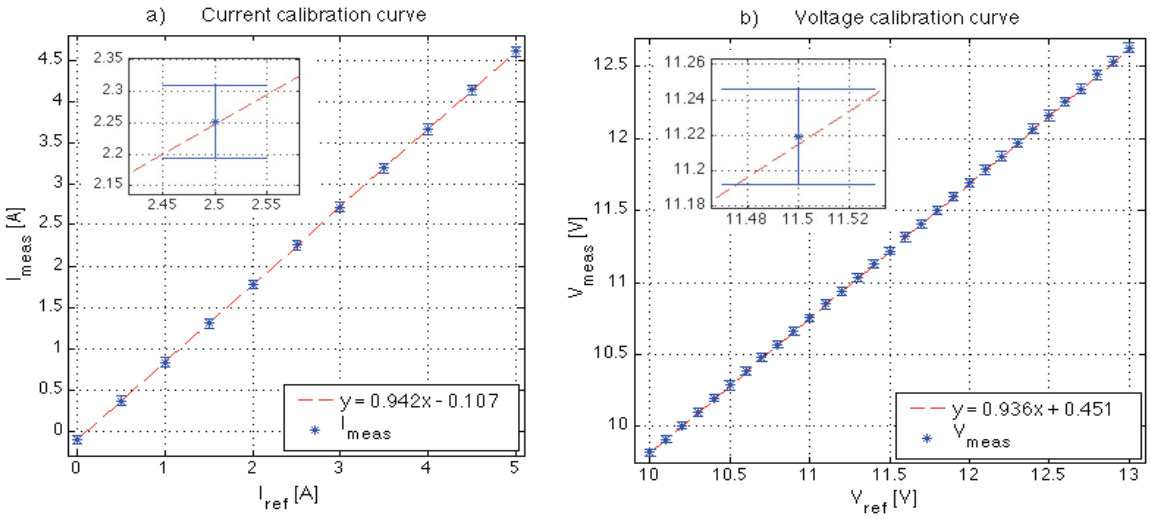

Figure 4. Current and voltage calibration curves: I_{meas} and V_{meas} represent the measured values from the power measurement system, I_{ref} and V_{ref} are the measured values from the Fluke 5500A calibrator.

C. Power measurement system calibration

The power measurement system has been calibrated by using the Fluke 5500A calibrator, which generates a reference value for the voltage and the current measures [10]. For the dc voltage range of 0 V to 33 V, the absolute uncertainty is around 1.7 mV with a resolution of 10 µV [10]. For the dc current range of 0 A to 11 A, the absolute uncertainty is around 6.9 mA with a resolution of 100 µA [10].

According to [8] the ADC resolution is around of 5 mV, therefore, by using the sensitivity value of 185 mV/A, the resolution for the current is 27 mA. For these reasons, Fluke 5500A has been chosen as reference calibration instrument for the voltage and the current measurements. The developed power measurement system has been calibrated separately for the current and voltage measurements according to the guideline [11]. The current values from the Fluke calibrator have been generated in the measurement range of the ACS714, 0 A to 5 A, with a step of 0.5 A. For each reference current value, 1000 measurements have been carried out using the realized system. Figure 4a shows the average and standard deviation of the measured values versus the reference ones. The measured results presented a systematic error which was compensated by using the first order regression curve y = 0.942x – 0.107 [A], also plotted in Fig.4a. The voltage values from the reference instrument have been generated in the measurement range of 10 ÷ 13 V, according to the tolerance value of ±5% allowed on the 12V line of the ATX Power Supply [12]. For each reference voltage value, 1000 measurements have been carried out using the realized system. Figure 4b shows the average and standard deviation of the measured values versus the reference ones. The measured results presented a systematic error which was compensated by using the first order regression curve y = 0.936x + 0.451 [V], also plotted in Fig.4b. For each point of the calibration curve has been evaluated the standard deviation both for current and voltage values. Once compensated for the systematic effects, the proposed power measurement system guarantees a current measurement range between 0 ÷ 5 A with a worst case experimental standard deviation of 33 mA and a voltage measurement range between 10 ÷ 13 V with a worst case experimental standard deviation of 17 mV. By neglecting the load effect of the measurement system and the resulting impairment and multiplying the current values by 2, it is possible to measure the power consumption up 130 W, having an estimated combined standard uncertainty of 0.84 W, according to [13]. Using a coverage factor of 2, it is possible to estimate an expanded uncertainty value of 1.68 W.

III. Description of the benchmarks

In [5], the authors have shown that the power consumption of several processors is heavily dependent on the characteristics of the software being executed. The long-term objective of the research described in this paper is to correlate the power consumption to the microscopic internal factors related to the execution of a given code (cache misses, mispredicted branches, number of active threads, page faults, use of floating point unit and of SIMD instructions, just to mention a few). For the time being, the measurements have been performed on a single CPU type, running freely available benchmark codes. To this aim, the benchmark suite called SiSoftware Sandra [14] has been chosen. This is a popular software able to stress various components of the CPU-memory system, and so the expectation was to observe a very different behavior as far as power consumption is concerned among the different tests making up the suite.

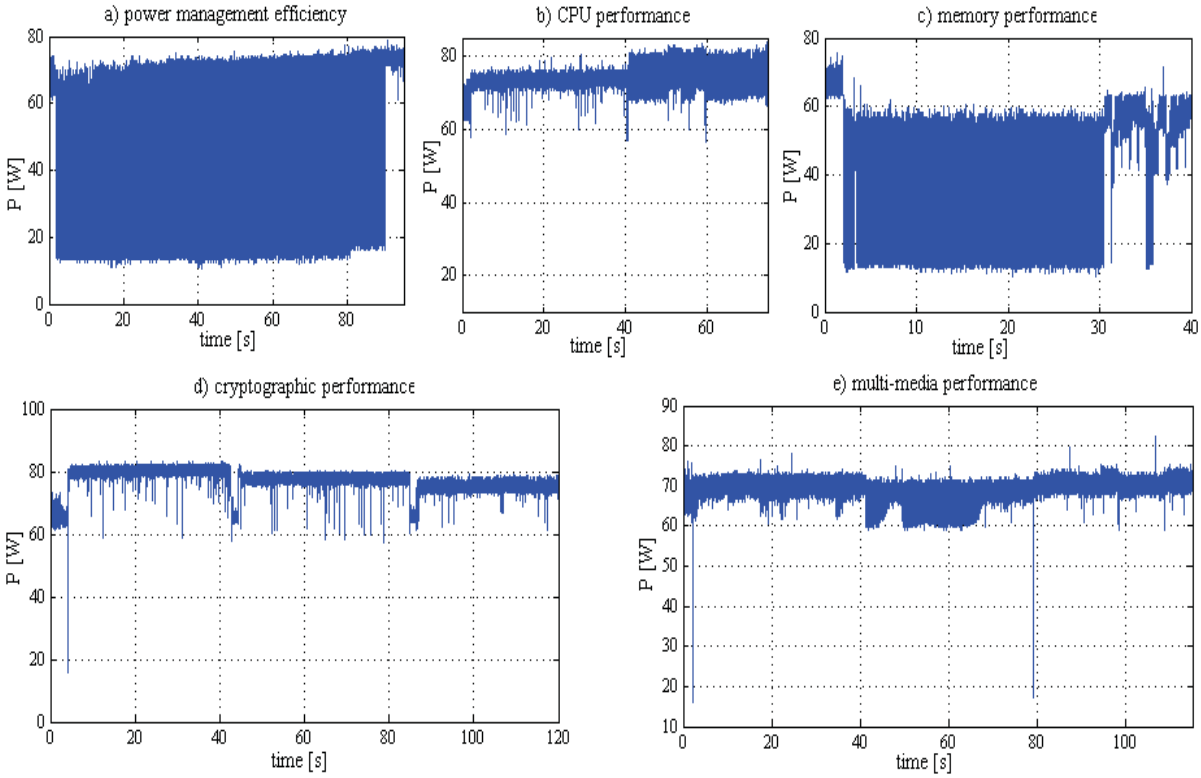

Figure 5. Measurements of the instantaneous power consumption during the execution of the benchmarks.

More detailed, SiSoftware Sandra is a 32 and 64 bit Windows based system analyzer, which includes the following benchmarking modules: (i) power management efficiency, (ii) CPU performance, (iii) memory performance, (iv) cryptographic performance, and (v) multi-media performance. The benchmark for the evaluation of the power management efficiency is designed to perform the processor benchmarking during execution time, when the processor is into the highest-performance execution state. The SiSoftware Sandra provides the Arithmetic and Logic Unit (ALU) and the Floating Point Unit (FPU) benchmark performance in terms of Millions of Instructions Per Second (MIPS) and Millions of FLoating-point OPeration per Second (MFLOPS). The benchmark of CPU performance has been used for the evaluation of the Giga Operations Per Second (GOPS) parameter. In order to measure the memory bandwidth in GB/s, the benchmark of memory performance has been executed. The benchmark of cryptographic performance was tested in order to measure the number of data that can be encrypted or decrypted per second, in MB/s, for the following algorithms: the Advanced Encryption Standards AES256-AES128 and the Secure Hash Algorithms for signing the SHA256-SHA1. Finally, the benchmark of CPU multi-media generates a picture (640x480) of the Mandelbrot fractal by using 255 iterations for each data pixel in 32 colors and provides the pixel rate of the algorithm.

IV. Results and discussions

The instantaneous power measurements on the WUT for the above benchmarks are presented in Figure 5: *(a)* power management efficiency benchmark, *(b)* CPU performance benchmark, *(c)* memory performance benchmark, *(d)* cryptographic performance benchmark, and *(e)* multi-media performance benchmark. For the mentioned cases, the power consumption measurements have been performed when the benchmark goes from the idle to the execution states.

After running the power management efficiency benchmark, it has been obtained a value of 617 executed MIPS. By using the target TDP value of 76.44 W, it is possible to estimate the power efficiency factor of 8.07 MIPS/W. It can be seen from the Fig.5a that during the benchmark execution, the power consumption changes continuously between a minimum value of 10 W and a maximum value of 80 W. In this case, by using this maximum value, it has been evaluated a power efficiency factor of 7 MIPS/W. The CPU performance benchmark provided 6.77 GOPS. The observed power consumption assumes a minimum value of 56 W and a maximum value of 84 W (see Fig.5b). The memory performance benchmark estimated the memory bandwidth of 3 GB/s. From Fig.5c, it can be said that the power consumption at the beginning and at the end of the benchmark execution is higher than power consumption in the middle. The cryptographic performance benchmark provided

a value of 42 MB/s for evaluating the number of data that can be encrypted or decrypted per second. It can be said that the power consumption decreases with the execution time from an initial value of 82 W to a final value of 79 W (see Fig.5d). The multi-media performance benchmark provided a rate of 6.76 MPixel/s. In this case, the power consumption is fairly constant over time (see Fig.5e).
The highest average power consumption (77 W) is measured for the benchmark *d* (cryptographic performance). Even if it has not been possible to examine the benchmark source code (SiSoftware Sandra is free, but not open-source), this is likely to be due to the use of streaming SIMD instruction for cryptographic purposes. Power consumption is slightly lower for the benchmark involving the intensive use of floating point operations (benchmarks *b* and *e,* CPU and multimedia performance, respectively). In the case *c* (memory performance), as mostly data movement operations to/from memory are performed, the average power consumption is much lower (31 W). As was expected, there is a significant difference in power consumption depending on the type of code being run. Compute-intensive codes can double the "normal" power consumption of the Pentium 4 under test. We believe that the punctual knowledge of the correlation between code actions and power consumption that is the objective of our future research can enable an energy-aware style of program writing. This will be of particular interest in the case of devices bound to low power consumption, as mobile systems running on batteries.

V. Conclusions and future works

A power measurement system designed for the measuring the instantaneous power consumption of a processor has been proposed. The architectural overview of the measurement system and the calibration procedure has been described. The power consumption measurements of a workstation computer for five benchmark applications, available in the SiSoftware Sandra suite, have been presented. Further work is directed to improve the measurement system, in order to measure the power consumption directly on the power supply pins of the processor, to correlate the power consumption with the processed events by processor, and to develop a mathematical model of power consumption to be applied/tested either for mobile applications.

Acknowledgment

The authors wish to thanks Prof. Pasquale Daponte and Assistant Prof. Luca De Vito for their helpful suggestions during all the phases of the present work.

References

[1] Jan et al., "A 22nm SoC platform technology featuring 3-D tri-gate and high-k/metal gate, optimized for ultra low power, high performance and high density SoC applications". Proc. of IEEE Int. Electron Devices Meeting (IEDM), San Francisco, USA, 10-13 Dec. 2012, pp.3.1.1-3.1.4.
[2] V.Zyuban, D.Brooks, V.Srinivasan, M.Gschwind, P. Bose, P.N.Strenski, P.G.Emma, "Integrated analysis of power and performance for pipelined microprocessors". IEEE Trans. on Computers, vol.53, Aug. 2004, No.8, pp.1004-1016.
[3] M.Moudgill, P.Bose, J.H.Moreno, "Validation of Turandot, a fast processor model for microarchitecture exploration". Proc. of IEEE Int. Conf. on Performance, Computing and Communications, Scottsdale, USA, 10-12 Feb. 1999, pp.451-457.
[4] S.Huck, "Measuring Processor Power". White Paper of Intel Xeon processor. Available online: http://www.intel.eu.
[5] P.Stanley-Marbell, V.C.Cabezas, "Performance, power, and thermal analysis of low-power processors for scale-out systems". Proc. of IEEE Int. Symposium on Parallel and Distributed Processing Workshops and Phd Forum (IPDPSW), Shanghai, 16-20 May 2011, pp.863-870.
[6] H.Esmaeilzadeh, T.Cao, Xi.Yang, S.M.Blackburn and K.S.McKinley, "Looking back on the language and hardware revolutions: measured power, performance and scaling". Proc. of ASPLOS XVI, Newport Beach, USA, 5-11 March 2011, pp.319-332.
[7] " "P4C800-E Deluxe". User guide, http://www.hp.com/ctg/Manual/c00054680.pdf.
[8] "PIC18F2455/2550/4455/3550 Data Sheet". Microchip, http://www.microchip.com.
[9] "Allegro ACS714 Data Sheet". Allegro, http://www.allegromicro.com.
[10] "Fluke 5500A Multi-Product Calibrator". Operator manual, http://www.fluke.com.
[11] "Servizio Nazionale di Taratura DT-05-DT rev.00". Available online: http://www.sit-italia.it/.
[12] "ATX Specification". Available online: http://pcsupport.about.com/od/insidethepc/a/power-supply-voltage-tolerance.htm.
[13] Guide to the Expression of Uncertainty in Measurement, GUM. Available online: http://www.bipm.org/en/publications/guides/gum.html.
[14] SiSoftware Sandra 2013". http://www.sisoftware.co.uk, April 2013.