Microprocessors and Microsystems 48 (2017) 3–10

Contents lists available at ScienceDirect

Microprocessors and Microsystems

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

Fine-grained monitoring for self-aware embedded systems

Mohamad Najem, Mohamad El Ahmad*, Pascal Benoit, Gilles Sassatelli, Lionel Torres

LIRMM, UMR CNRS 5506, University of Montpellier, France

ARTICLE INFO

Article history:
Received 22 February 2016
Revised 2 June 2016
Accepted 16 September 2016
Available online 23 September 2016

Keywords:
Power monitoring
Modeling
Performance events
Selection
Neural network

ABSTRACT

Dynamic Thermal and Power Management methods highly depend on the quality of the monitoring, which needs to provide estimations of the system's state. This can be achieved with a set of performance counters that can be configured to track logical events at different levels. Although this problem has been addressed in the literature, recently developed highly reactive adaptation techniques require faster, more accurate and more robust estimations methods. A systematic approach (PESel) is proposed for the selection of the relevant performance events from the local, shared and system resources. We investigate an implementation of a neural network based estimation technique which provides better results compared to related works. Our approach is robust to external temperature variations and takes into account dynamic scaling of the operating frequency. It achieves 96% accuracy with a temporal resolution of 100 ms, with negligible performance/energy overheads (less than 1%).

1. Introduction

Achieving high energy efficiency is a major challenge in the design of integrated circuits, which is confronted with the problem of delivering high performance with a limited power budget. Therefore, modern day microprocessors provide dynamic thermal and power management (DTPM) techniques to address this challenge. Similarly to the definition in [1], a "self-aware" approach assumes that the adaptation is not governed by an external process but initiated by the system itself to improve its own performance, which requires an integrated monitoring structure allowing the self-observation of the system's state. Several DTPM techniques use the information of the power consumed by different units of the system, to either adapt the system behavior (reactive techniques [2]) or to predict future undesired states such as critical chip temperature (proactive techniques [3]). In all cases, effective management of power and temperature depends critically on the monitoring method, which should provide high-frequency, accurate and fine grain estimations.

Nowadays, systems integrate several cores grouped together into clusters, such as ARM Big.LITTLE from Samsung and Intel Xeon Sandy Bridge-EP [4]. Also, these systems integrate a specific unit connected to several event counters for debugging and performance profiling. This unit is referred to as Performance Monitoring Unit (PMU) in ARM based processors and Performance Monitoring Counters (PMCs) in Intel based processors. In the rest of this paper, we choose the notation of ARM to designate it. As an illustration, Fig. 1 shows an example of a generic cluster that contains several cores sharing some peripherals and components, including local interconnect, L2 cache and the PMU. Events that can be monitored in such systems can be decomposed into three main categories: (i) local events occurring at the hardware-level inside each core (e.g. L1 instruction misses), (ii) shared events occurring at the hardware-level on the shared resource inside the cluster (e.g. L2 cache accesses), and (iii) events from the operating system (e.g. a task migration). The PMU has a dedicated hardware counter to monitor the occurrence of both local and shared events, while the system event counters are implemented in a software layer managed by a firmware. PMUs integrate a few numbers of hardware counters, less than the number of available events, which can be configured to monitor a subset of events. For instance, the PMU in ARMv7 Cortex A class processors has 6 hardware counters to monitor 62 local events, and 2 counters for 15 shared events while the PMC in Xeon processors from Intel has only 4 counters for 160 of both local and shared events.

In this paper, we address the self-awareness in embedded systems, especially for the monitoring of the power consumption. The goal is to estimate at run-time the overall consumption, which is due to both switching (dynamic power) and leakage (static power). The overall power consumption depends on power supply voltage, circuit frequency, internal activity, but also temperature and transistor characteristics. The internal activity is one of the parameters that dynamically varies with the applications for a given functional mode (power supply and frequency), while the operating temper-

* Corresponding author.
E-mail addresses: najem.mhd@gmail.com (M. Najem), mohamad.elahmad@lirmm.fr, mohamad-elahmad@hotmail.com (M.E. Ahmad), pascal.benoit@lirmm.fr (P. Benoit), gilles.sassatelli@lirmm.fr (G. Sassatelli), lionel.torres@lirmm.fr (L. Torres).

http://dx.doi.org/10.1016/j.micpro.2016.09.004
0141-9331/

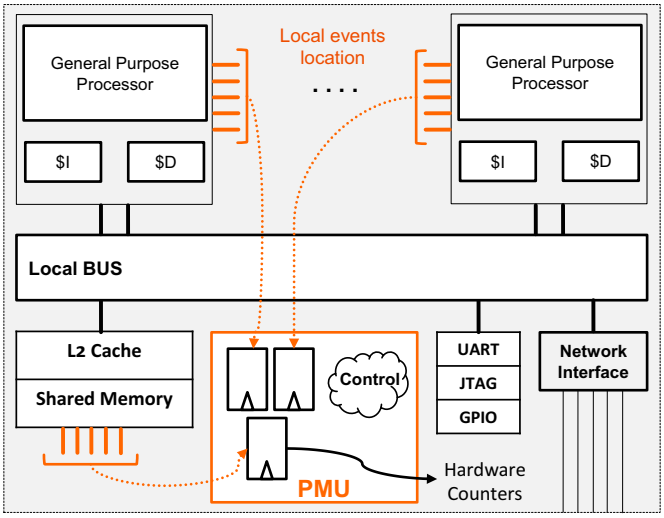

Fig. 1. Example of a cluster, in which a dedicated performance events monitoring unit.

ature depends directly on the environment temperature. The main challenge is to find the best subset of events that appraises the global internal activity, and to build a model that takes into account the other parameters. For this purpose, we investigate in this paper power models robust to external temperature variations. We use the information from the PMU to estimate the activity, an external temperature sensor, and the information about cores frequencies to accurately track any variation in the power consumption. There are thousands of events that can be monitored, each of which may generate high frequency streams of activity data. Handling such an enormous amount of data is particularly challenging and requires suitable techniques. For this purpose, we developed the algorithm *PESel* (Performance Events Selection) inspired from the features selection methods from Data Mining to select relevant events for power modeling purposes. Furthermore, a linear and neural network models are compared for tracking the variation of the power at a high time resolution.

The remainder of this paper is organized as follows. In Section II the limitations of most relevant power estimation and event selection techniques are discussed to motivate this work. Section III and IV present the proposed selection method and power models. Section V describes the experiments and Section VI draws conclusions from those results and discusses future lines of research.

2. Background and related works

The total power consumption, which is that consumed by the device during a specific time interval, can be expressed as:

$$P_{total} = P_{dyn} + P_{stat} = (P_{sw} + P_{sc}) + P_{stat} = \left(\alpha \cdot C \cdot V_{dd}^2 \cdot f + V_{dd} \cdot I_{sc}\right) + V_{dd} \cdot I_{leakage} \quad (1)$$

Where α is the activity factor, C the switching capacitance, V_{dd} the power supply, and f the operating frequency. I_{sc} is the short-circuit current, and occurs when pull-up and pull-down networks are conducting simultaneously. It has a low impact on the average power consumption over a time interval, and is usually neglected.

The dynamic part dominates during active mode of the system, whereas static part offsets total power as long as the system is powered on. Static power increases exponentially with the chip temperature (T_{chip}). This dependence can result in an overall worsening, because temperature is also a function of power consumption. With this assumption static power (P_{stat}) can be expressed as:

$$P_{stat} = P_0 \cdot e^{-k_\theta / T_{chip}} \quad (2)$$

Where P_0 and k_θ depends on supply voltage and thermal characteristics of the chip. Consequently, the updated equation for the power dissipation can be expressed as follows:

$$P_{total} = \alpha \cdot C \cdot V_{dd}^2 \cdot f + P_0 \cdot e^{-k_\theta / T_{chip}} \quad (3)$$

According to [5], the external temperature T_{ext} and T_{chip} (operating temperature) may be wrapped into the same equation by using the equivalent RC circuit for modeling the temperature, and can be expressed as follows:

$$T_{chip} = T_{ext} + R_\theta \cdot P_{total} \quad (4)$$

Where R_θ is the equivalent thermal resistance of the package (°C/W). Since the ambient temperature has a direct impact on T_{chip}, it is required to account for this parameter in the power models for tracking the power consumption.

Many research efforts have focused on how power consumption can be tracked considering only the internal activity (α), which is usually appraised by a set of performance events collected at run-time using the PMU. In [6], authors have used a simple linear combination of few performance events (e.g. the number of instructions executed) to estimate the total energy. The selection of events was manually done based on a previous knowledge about the system behavior. This approach may have a good accuracy in some cases, but the selection of events is very application specific. The selection of the relevant events for the power modeling was also addressed in the literature. Authors in [7] have selected the performance events using the Pearson correlation criterion, while in [8],[9] it was based on the Spearman's rank correlation. These statistical criteria are quite limited and do not consider the interaction between the selected events. However, perfectly correlated events can be redundant in the sense that no additional information is gained by keeping both; conversely two events that are separately poorly correlated with power can prove useful when exploited together for power modeling [10]. A projection method based on the principle component analysis (PCA) was proposed in [11] that takes into account the relation between the elements before the selection of performance events for the Dell PowerEdge Opteron processor. However, events are selected for each application separately from the PB-3.3-OM benchmark, and are used to train a simple linear model with a high data sampling period (300 ms). This is an interesting approach, but requires the reconfiguration of events at run-time for each executed application, and does not consider the simultaneous execution of the applications in multiprocessor systems. Also, due to complex logical interactions, simple linear models might fail in accurately tracking power.

Some of the research papers in the domain assume a quasi-constant average activity, and focus more on the frequency and chip temperature to estimate power consumption. In [12], authors assume constant average internal activity (α) by application while modeling the power using the Eq. (3) for each of the cluster in the ARM big.LITLLE platform. Furthermore, authors in [13] neglect the temperature representing the power as a polynomial equation of the couple frequency and supply voltage. These models may have a good accuracy in some cases especially for the estimation of the energy budgets. However, the internal activity is one of the main parameters that significantly affects the power consumption and must be represented in order to accurately observe fine-grain power variations. For power modeling, some usual techniques that have been proposed are shown in Table 3.

Our objective is to address fine-grained power monitoring, taking into account the internal activities but also the whole set of parameters that may change the consumption.