Simulation Modelling Practice and Theory 36 (2013) 33–43

Contents lists available at SciVerse ScienceDirect

Simulation Modelling Practice and Theory

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

Performance evaluation and energy consumption of a real-time heterogeneous grid system using DVS and DPM

George Terzopoulos *, Helen Karatza [1]

Department of Informatics, Aristotle University of Thessaloniki, Greece

ARTICLE INFO

Article history:
Received 4 March 2013
Received in revised form 24 April 2013
Accepted 25 April 2013

Keywords:
Performance
Simulation
Grid
Power-saving
DVS
DPM

ABSTRACT

Energy consumption of large scale systems has been severely studied due to economic and ecological reasons. This paper studies energy gains that come from the application of two popular energy saving techniques, Dynamic Voltage Scaling (DVS) and Dynamic Power Management (DPM), in a real-time 2-level heterogeneous grid system. While these techniques generally work in a competitive way, we show that under certain circumstances they can work together and achieve greater savings when they are both applied at the processor level. A simulation model is used to evaluate the performance of the system. Experimental results show encouraging energy savings up to 46% and minimum performance degradation when both energy saving techniques are applied.

1. Introduction

Energy consumption of computer systems has recently gained significant attention due to the following reasons:

- A system that consumes less energy, generates less heat and therefore has a longer lifetime. According to a formula based on the Arrhenius Law, component life expectancy decreases 50% for every 10 °C temperature increase. Thus, reducing a component's operating temperature by the same amount (consuming less energy), doubles its life expectancy [1].
- In battery operated systems, a crucial objective is to work under the minimum power consumption and provide a desired level of performance at the same time. Portable devices need to preserve energy in order to increase autonomy.
- The heat in large scale computers requires expensive cooling technologies. In order for large companies to reduce cooling cost they consider building large server farms in cold places. Facebook is to build a multi-million server farm on the edge of the Arctic Circle in Northern Sweden. The center will be managing all of Facebook's traffic in Europe, serving more than 180 million users.
- Electricity cost of the computing hardware is very high [2]. Estimations show that computing center servers currently consume about 0.5% of the total electricity consumption in the world, and the percentage will be quadruple by 2020 if the current demand keeps rising [3].
- Environmental reasons. The global information and communications technology (ICT) industry accounts for approximately 2% of global carbon dioxide (CO_2) emissions, according to an estimate by Gartner Inc. [4].

* Corresponding author. Tel.: +30 6936792489.
E-mail addresses: gterzopo@csd.auth.gr (G. Terzopoulos), karatza@csd.auth.gr (H. Karatza).
[1] Tel.: +30 2310997974.

1569-190X/$ - see front matter
http://dx.doi.org/10.1016/j.simpat.2013.04.006

Research has focused on reducing processors' power consumption in computing systems. There are currently two popular power minimization techniques: Dynamic Power Management (DPM) [5] and Dynamic Voltage Scaling (DVS) [6].

DPM puts the idle components of the system into low-power sleep states whenever this is possible. Usually, components are turned off after a fixed amount of idling time. More advanced methods try to predict the expected idling time using past data in order to estimate the future status of a system. At the processor level, without the use of DPM, every processor is in active or idle state during runtime. In a system where the load is low, a lot of processors will be in idle state and consume energy without any system product. Excess processors could be switched into sleep state. Modern processors support several sleep power modes called C-States, starting from C0. The higher the C number is, the deeper is the CPU sleep mode and while CPU consumes less energy, it needs more time to wake-up and be 100% operational. Thus, energy and time overheads should be taken into account caused by sleep/wake power transitions.

DVS trades off performance for power consumption by lowering the operating voltage/frequency. It dynamically scales the processor's core voltage depending on the computational demand of the system and due to the quadratic relationship between voltage and dynamic power, DVS achieves significant power savings. DVS is widely supported in new processors. Both AMD (AMD Cool'n'Quiet, AMD PowerNow!) and Intel (Intel SpeedStep) processors support DVS [7].

2. Related work

In real-time systems, each task has a certain deadline associated with it and the task has to finish its execution before that deadline. When the processor's speed is reduced with DVS, a scheduling algorithm must guarantee that deadlines can still be met at the adjusted speeds. DVS schemes are exploited in real-time systems where processors adjust their operating frequency in order to execute tasks within their deadline [8–10]. In [11], authors study energy-efficient elastic scheduling for independent tasks on DVS-enabled heterogeneous computing systems. Their objective is to make the best trade-off between energy efficiency and users' expectation.

Slowing down the processor's speed when DVS is applied stretches task execution time and this leads to reduction of the energy consumed by the processor. In a system where a processor interacts with other devices, the total energy consumption may be increased. To address this problem, authors in [12] consider the system-level energy consumption and not only the power consumption of processors and propose techniques to reduce the overall energy consumption of the system.

There are several studies that suggest the application of both DPM and DVS techniques to maximize energy savings. In most studies, DVS is used to reduce processors' energy consumption and DPM is used in order to shut down other peripheral devices [13–15]. In [16], a combined DVS–DPM approach at the network level in a distributed real-time embedded system consisting of a set of wireless nodes is proposed. Processors are considered to be ideal. An ideal processor can operate at any speed, while a non-ideal processor can operate only at discrete speeds. In [17], the interplay between DVS and DPM for a real-time application that uses several devices is studied taking into account device transitions to minimize the system-wide energy consumption. DVS is applied to the processor and DPM is applied to other devices. In [18], authors propose power management schemes for heterogeneous clusters under QoS constraints. Cluster's servers are shut down when they are not needed although power consumption of inactive servers is considered to be zero. In [19], a real-time system with a DVS-enabled processor and a fixed number of offchip devices used by the tasks during their executions is studied. In this approach, DPM is applied at the offchip devices and not at the processors. In [20], resource allocation policies in a heterogeneous cluster taking into account energy consumption are studied without the application of DVS. Their system consists of eight energy efficient processors with low performance and eight high performance processors with high energy consumption. In a previous research [21], we proposed the application of a DVS mechanism in a real-time heterogeneous 2-level grid. The results were very promising. We achieved energy savings 14–35% without significant performance degradation.

In this paper we study the performance and the energy efficiency of a real-time distributed system with four heterogeneous clusters with DVS processors that can adjust their operating frequency depending on the workload. In addition to our previous work, we apply both DVS and DPM at the processor level. Our goal is to minimize the energy consumption of the processors without sacrificing system's performance. With the application of power saving mechanisms, we achieved energy consumption reduction 15–46%, depending on the workload of the system.

With respect to the above-mentioned related work, our proposed approach is novel in that we apply both DVS and DPM at the processor level, in order to preserve energy and maintain performance at high levels. In our approach, DVS is not applied constantly but under certain circumstances depending on the workload. System's processors support a limited number of discrete frequency levels and are not considered to be ideal, in addition to aforementioned studies. Since processors are often the dominant consumers of power in servers we focus only on processors' energy consumption. Our DPM power-saving technique is applied only when system's load is relatively low and the power consumption of processors that are in sleep state is taken into account. In addition, useful conclusions are made about the impact of load variability to the system's energy consumption and performance, due to fact that the system is evaluated when there is for low and high variability in the arrival rate. The interaction of two power-saving techniques, DVS and DPM, and the conditions under which energy savings from both techniques are maximized, are also evaluated. Hard real-time tasks submitted to our system have different QoS requirements, thus different deadlines, and this leads to variations in deadline miss ratio for every type of task. In a real-time system with tasks that need to complete execution within a deadline, a combination of scheduling policies and power saving mechanisms is needed in order to have high performance and low energy consumption at the same time. To the best of our knowl-