Electrical Power and Energy Systems 78 (2016) 499–512

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

Joint optimization of energy efficiency and system reliability for precedence constrained tasks in heterogeneous systems

Longxin Zhang [a], Kenli Li [a,*], Keqin Li [a,b], Yuming Xu [a]

[a] College of Computer Science and Electronic Engineering, National Supercomputing Center in Changsha, Hunan University, Changsha 410082, China
[b] Department of Computer Science, State University of New York, New Paltz, NY 12561, USA

ARTICLE INFO

Article history:
Received 6 August 2014
Received in revised form 23 September 2015
Accepted 25 November 2015
Available online 23 December 2015

Keywords:
Energy management
Parallel application
Reliability enhancement
Scheduling algorithm

ABSTRACT

Voltage scaling is a fundamental technique in the energy efficient computing field. Recent studies tackling this topic show degraded system reliability as frequency scales. To address this conflict, the subject of reliability aware power management (RAPM) has been extensively explored and is still under investigation. Heterogeneous Computing Systems (HCS) provide high performance potential which attracts researchers to consider these systems. Unfortunately, the existing scheduling algorithms for precedence constrained tasks with shared deadline in HCS do not adequately consider reliability conservation. In this study, we design joint optimization schemes of energy efficiency and system reliability for directed acyclic graph (DAG) by adopting the shared recovery technique, which can achieve high system reliability and noticeable energy preservation. To the best of our knowledge, this is the first time to address the problem in HCS. The extensive comparative evaluation studies for both randomly generated and some real-world applications graphs show that our scheduling algorithms are compelling in terms of enhancement of both system reliability and energy saving.

Introduction

During the past decades, resource management and task scheduling on computer system have been well addressed. With the advent of multiprocessor and multi-core, energy management has become one of the hottest areas due to the dramatically increasing power consumption of modern heterogeneous computational systems, which consist of various set of resources. In reducing energy consumption field, one of the recent main challenges is to accomplish speed up on parallel application with regard to directed acyclic graphs (DAG). It's generally considered that task scheduling problem is an NP-complete problem [1].

The new generation of processors provides a substantial performance boost meanwhile, the energy consumption is not a well addressed problem. Although a lot of effort is being spent on saving energy of processor, the efficient and effective method has yet to be developed. The emergence of severe ecological, economic and technical issues are caused by the increasing energy consumption. Hence, it is not very hard to image the size of adverse environmental footprint left by HCS. With the growing advocacy of green computing systems, the issue of energy efficiency has recently attracted extensive research. Hardware technologies [2], energy-efficient monitors, low-power microprocessors, processor contains multiple processing element cores and a selective-associative cache memory, are employed to deal with the energy consumption problems. Comprehensive surveys can be referred to references [2–4].

Dynamic voltage frequency scaling (DVFS) is a fundamental energy management technique for modern computing systems. It has been proven that the dynamic power consumption is a strictly convex function of the CPU speed. Reducing clock speed or supply voltage of processor at active state to achieve energy saving are examples for energy consumption reduction. Various energy efficient scheduling and resource management strategies have developed for sustainable computing via taking advantage of DVFS technique. Nevertheless, the excellent schemes for these scope of applications are limited to unique processor systems [5], battery based embedded systems [6] or homogeneous computing systems [7–9].

Generally speaking, task scheduling is to select a proper processor for each task which complies with the precedence rules. In the meanwhile, the main goal is to minimize scheduling length. One relative research is the famous Dynamic-Level Scheduling (DLS) algorithm [10] for homogeneous systems. During the past few years, some list scheduling algorithms are proposed for heterogeneous

* Corresponding author.
E-mail addresses: longxinzhang@hnu.edu.cn (L. Zhang), lkl@hnu.edu.cn (K. Li), lik@newpaltz.edu (K. Li), xxl1205@163.com (Y. Xu).

http://dx.doi.org/10.1016/j.ijepes.2015.11.102
0142-0615/

systems, such as Constrained earliest finish time (CEFT) algorithm [11], Critical-Path-On-a-Processor (CPOP) algorithm [12], and heterogeneous earliest finish time (HEFT) algorithm [12]. HEFT and CPOP are famous for its low-complexity and performance-effective capability. CEFT is capable of accomplishing tasks earlier without sacrificing too much time complexity.

Another important challenge is to avoid communication between processors while transferring data for task nodes with dependence. An advance technique, named node duplication, is considered as an effective solution to the stated problem. Based on this, some recently reported researches, i.e., HCPFD [13] and HLD [14], are proposed to achieve better performance. The main idea of these works is to duplicate the immediate parent of each task if possible and necessary so as to diminish unnecessary communication between parent and child nodes, which are allocated to the distinct processors. The extensive comparisons, in [13,14], demonstrate that algorithms with idea of duplication significantly surpass other strategies, such as DLS [10] and HEFT [12].

Recently, Zhang et al. [15] presented algorithms to enhance task reliability when a dynamic voltage scaling technique is applied to achieve low power consumption in heterogeneous computing systems. Xu et al. [16] proposed a hybrid chemical reaction optimization scheme for task scheduling on heterogeneous computing systems. Huang et al. [17] developed novel heuristic speculative execution strategies in heterogeneous distributed environments. Yan et al. [18] developed an intelligent particle swarm optimization (IPSO) algorithm for short-term traffic flow predictors to tackle time-invariant assumptions in the on-road sensor systems.

Unfortunately, most of these approaches are on the basis of simple system models. It is inadequate in two respects. On one hand, the employed system model does not precisely reflect the real parallel computation system. On the other hand, the assumption that each task will always execute successfully is not occur with real application. For various reasons, transient faults arise more frequently than permanent faults. In particular, transient fault is a small probability event but inevitably occurs in the real situation. The side effect of transient fault is that, it may cause seriously adverse hurt on the running application, sometimes it even leads to fatal errors. Low-power consumption and high-system reliability, availability, and utility are major concerns of modern high-performance computing system development. There are a few studies focusing on providing fault tolerance capabilities to embedded systems [19]. Studies in [20,21] explore the interplay between reliability and energy consumption. However, they are exclusively confined to the field of real-time system. With regard to HCS, Tang et al. [22] developed a reliability-driven scheduling architecture and devised a reliability-aware scheduling scheme for precedence constrained tasks. Lee and Zomaya [23] proposed two energy-conscious scheduling algorithms which effectively balance the quality of schedules and energy consumption with DVFS technique. Li [24] analyzed the performance of heuristic power allocation and scheduling algorithms for precedence constrained sequential tasks. Notwithstanding, none of them incorporates energy consumption and reliability together. In most cases, the smaller scheduling length is not the most important objective. The co-management of system reliability and energy saving has drawn researchers' attention only very recently [19,25,26]. Zhao et al. [19] adopted recovery technique to reexecute fault application for tasks with dependence, but it is only applied to unique core.

In this study, we develop energy efficient and reliability conservation for precedence constrained tasks with shared deadline in heterogeneous system. It's also a combinatorial optimization problem. To the best of our knowledge, this is the first time the problem is dealt with in HCS. Two kinds of techniques, dynamical voltage scaling and shared recovery are employed. Our scheduling problems comprise five nontrivial subproblems, namely, precedence constraining, deadline constraining, energy conservation, reliability maximizing and task recovery.

- *Precedence Constraining*. Compared to independent task set, tasks with precedence constraints make devise and analysis of heuristic scheduling algorithms particularly complicated.
- *Deadline Constraining*. The entire task set share a common deadline.
- *Energy Conserving*. Tasks should be supplied with appropriate powers and energy efficient execution speeds, such that the schedule length is modest and the energy consumption is minimal.
- *Reliability Maximizing*. Tasks should be run at a relatively high speed without exceeding the maximum frequency of processor, such that the system reliability can achieve an optimal value.
- *Task Recovery*. After finish each task in the prior queue, error detection would be started up. Once a fault occurs, then the error will be recovered at the maximum frequency.

The above subproblems should be solved efficiently so that heuristic algorithms with overall fine performance can be explored. By adopting DVFS technique, three algorithms will be presented in this paper: Shared Recovery for reliability aware of Heterogeneous Earliest Finish Time (SHRHEFT) algorithm, Shared recovery for Reliability aware of Critical Path On a Processor (SHRCPOP) algorithm and Shared Recovery for Energy Efficiency and system Reliability Maximization (SHREERM) algorithm. Each of these three algorithms comprises three phases. Task priority establishment phase builds a proper topological order for the application task. Frequency selection phase chooses an energy efficient frequency to execute each task. Processor assignment phase allocates the candidate task to a suitable processor. Finally, task recovery phase detects the transient fault and recovers the error task, so as to get higher system reliability with lower total energy consumption.

The paper proceeds as follows: Section 'Models' gives a briefly introduction of models used in this paper which include the system, power, energy and reliability. Based on reliability and energy model, SHRMEC algorithm and revised algorithms SHRHEFT and SHRCPOP are presented in Section 'The proposed algorithms'. Extensive experiment results are discussed in Section 'Experiment s and results'. Section 'Conclusion' summarizes the conclusion and shows the future direction of this work.

Models

In this section, we provide a brief background on the system, power, fault, and application models used in this paper.

System model

The system model used in this paper is composed of a set P that includes $|p|$ heterogeneous cores/processors. All the cores are encapsulated in a chip. Each core of them is available for DVFS technology; namely, each core can run at different speed (i.e., different supply voltage levels) in any time. For each processor p belongs to the set P, with $|f|$ available frequency levels (*AFLs*), follows a random and uniformly distribution among the four different sets of operation voltages/frequency. As clock frequency is switched, overheads take an inappreciable amount of time (e.g. 10–150 μs) [27,28], and this overheads are neglected in our study. The communications among inter-processor are supposed to perform at the same speed on all links without contentions. Our paper is based on the premise that the target system consists of a set of