Future Generation Computer Systems 29 (2013) 1661–1670

Contents lists available at SciVerse ScienceDirect

Future Generation Computer Systems

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

Energy-aware parallel task scheduling in a cluster

Lizhe Wang [a,b,*], Samee U. Khan [c,**], Dan Chen [a,**], Joanna Kołodziej [d], Rajiv Ranjan [e], Cheng-zhong Xu [f], Albert Zomaya [g]

[a] School of Computer Science, China University of Geosciences, PR China
[b] Center for Earth Observation and Digital Earth, Chinese Academy of Sciences, PR China
[c] Department of Electrical and Computer Engineering, North Dakota State University, United States
[d] Institute of Computer Science, Cracow University of Technology, Poland
[e] ICT Centre, CSIRO, Australia
[f] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, PR China
[g] School of Information Technologies, The University of Sydney, Australia

HIGHLIGHTS

- This paper develops formal models for parallel tasks and a power aware cluster.
- Power aware scheduling for parallel tasks based on list scheduling is proposed.
- System model is based on SLA for green parallel task scheduling.
- The proposed methodologies are thoroughly investigated through simulations.

ARTICLE INFO

Article history:
Received 23 September 2011
Received in revised form
19 November 2012
Accepted 25 February 2013
Available online 18 March 2013

Keywords:
Cluster computing
Green computing
Task scheduling

ABSTRACT

Reducing energy consumption for high end computing can bring various benefits such as reducing operating costs, increasing system reliability, and environmental respect. This paper aims to develop scheduling heuristics and to present application experience for reducing power consumption of parallel tasks in a cluster with the Dynamic Voltage Frequency Scaling (DVFS) technique. In this paper, formal models are presented for precedence-constrained parallel tasks, DVFS-enabled clusters, and energy consumption. This paper studies the slack time for non-critical jobs, extends their execution time and reduces the energy consumption without increasing the task's execution time as a whole. Additionally, Green Service Level Agreement is also considered in this paper. By increasing task execution time within an affordable limit, this paper develops scheduling heuristics to reduce energy consumption of a tasks execution and discusses the relationship between energy consumption and task execution time. Models and scheduling heuristics are examined with a simulation study. Test results justify the design and implementation of proposed energy aware scheduling heuristics in the paper.

1. Introduction

Nowadays, high end computing facilities can consume a very large amount of power, although they provide high performance computing solutions for scientific and engineering applications [1]. For example, operating a middle-sized data center (i.e., a university data center) demands 80 000 kW power [2]. It is estimated that computing resources consume around 0.5% of the world's total power usage [3], and if current demand continues, this is projected to quadruple by 2020. Energy consumption for high performance facilities thus contributes to a significant electric bill. Additionally, high power consumption in general results in higher cooling costs. Furthermore, to allow computing facilities to operate on high power for a long time will lead to a high temperature of computing systems, which further harms a system's reliability and availability. Therefore, reducing power consumption for high end computing becomes a critical research topic.

Modern processors are equipped with the Dynamic Voltage Frequency Scaling (DVFS) technique, which enables processors to be operated at multiple frequencies under different supply voltages. The DVFS technique thus gives opportunities to reduce the energy consumption of high performance computing by scaling processor supply voltages. Our research is devoted to developing

* Corresponding author at: School of Computer Science, China University of Geosciences, PR China.
** Corresponding authors.
E-mail addresses: Lizhe.Wang@gmail.com (L. Wang), samee.khan@ndsu.edu (S.U. Khan), Danjj43@gmail.com (D. Chen).

0167-739X/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.future.2013.02.010

scheduling heuristics which reduce energy consumption of parallel task execution by using the DVFS mechanism. A parallel task is a set of jobs with precedence constraints. Jobs in a parallel task may have some slack time for their execution due to their precedence constraints.

This paper makes a study of scheduling policies and application experiences to reduce power consumption of parallel tasks. Our first research issue is to minimize task execution time as well as reduce power consumption. The execution time of the non-critical jobs in a parallel task can be extended, thus giving an opportunity to scale down the supply voltages of processors. Based on the analysis of DVFS on non-critical jobs, we develop two power aware scheduling heuristics for parallel tasks, the Power Aware List-based Scheduling (PALS) algorithm and the Power Aware Task Clustering (PATC) algorithm.

Our second research objective is to make an study on energy and performance tradeoff for parallel task execution. The green Service Level Agreement (SLA) is introduced in this research. By negotiating with users via the green SLA, an energy-performance tradeoff algorithm is developed to reduce energy consumption with an affordable task execution time increase. We develop a simulation study on the proposed scheduling heuristics and make a performance evaluation.

We declare our contribution as follows:

- We develop formal models for parallel tasks and a power aware cluster and we also define the task scheduling issue.
- We develop two power scheduling heuristics for parallel tasks: the PALS and the PATC.
- We present the green SLA use scenarios and propose a new scheduling heuristics for energy aware parallel task scheduling, which makes a study of the tradeoff between the energy consumption and task execution time (performance).
- We build a simulation study and performance evaluation on the proposed heuristics. Test results justify our design and implementation of energy aware heuristics.

The rest of this paper is organized as follows. Section 2 introduces background and related work. Then Section 3 discusses the models for parallel tasks, DVFS and compute clusters and Section 4 formally define the research issue of energy aware parallel task scheduling. Section 5 applies the DVFS technique on non-critical jobs of parallel tasks, which is the basis of the PALS and the PATC. We describe the scheduling heuristics of the PATC and the PALS in Sections 6 and 7. Section 8 presents the Service Level Agreement with performance metrics of green computing and proposes the research issue of energy-performance tradeoff for parallel task scheduling. Section 9 then presents the scheduling algorithm for the research issue proposed in 9. The complexity analysis for the proposed algorithms are presented in Sections 10 and 11 describes a simulation study on the proposed scheduling heuristics. Finally this paper is summarized in Section 12.

2. Related work

This section discusses background and related work of task scheduling, DVFS, and power aware cluster computing.

2.1. Parallel task scheduling

Task scheduling techniques in parallel and distributed systems have been studied in great detail with the aim of making use of these systems efficiently.

Task scheduling algorithms are typically classified into two subcategories: static scheduling algorithms and dynamic scheduling algorithms. In static task scheduling algorithms, the task assignment to resources is determined before applications are executed. Information about task execution cost and communication time is supposed to be known at compilation time. Static task scheduling algorithms normally are non-preemptive—a task is always running on the resource to which it is assigned [4]. Dynamic task scheduling algorithms normally schedule tasks to resources in the runtime to achieving load balance among PEs. are based on the redistribution [5,6].

The list scheduling algorithm is the most popular algorithm in the static scheduling [7,8]. List based scheduling algorithms assign priorities to tasks and sort tasks into a list ordered in decreasing priority. Then tasks are scheduled based on the priorities. In this paper, we build a list based scheduling heuristic for parallel tasks—the PALS algorithm. The task execution information, such as task execution cost and communication cost, can be obtained by some profiling tools and compiler aides in advance.

The task graph clustering technique [9,10] is an effective static scheduling heuristic for scheduling parallel tasks. Given a task graph, "clustering" is the process of mapping task graph nodes onto labeled clusters. All tasks of the same cluster are executed in the same processor. In traditional task scheduling heuristics, the process of clustering tasks is an optimization of reducing the makespan of the scheduled graph. In this paper, we proposed the PATC algorithm, whose process of clustering tasks is guided by reducing the total power consumption of the scheduled graph.

2.2. Energy reduction via DVFS techniques

Dynamic voltage and frequency scaling (DVFS) has been proven to be a feasible solution to reduce processor power consumption [11,12]. By lowering processor clock frequency and supply voltage during some time slots, for example, idle or communication phases, large reductions in power consumption can be achieved with only modest performance losses. A DVFS-enabled cluster [1] is a compute cluster where compute nodes can run at multiple power/performance operating points. The DVFS techniques have been applied in the high performance computing fields, for example, in large data centers, to reduce power consumption and achieve high reliability and availability [13–15]. Popular DVFS-based software solutions for high end computing include:

- Scientific applications can be modeled with a Directed Acyclic Graph (DAG) and the critical path is identified for applications. Thus, it is possible to reduce energy consumption by leveling down the processor supply voltage during non-critical execution of tasks [16].
- Some work [17] builds online performance-driven runtime systems to automatically scale processor supply voltages.
- Some work applies DVFS during the communication phases of high performance computing, for example MPI [18,19].
- In addition to parallel applications, virtual machine scheduling can also use DVFS [1].

Our research in this paper falls into the first category: scheduling DAGs on multiple processors in a cluster with DVFS techniques.

2.3. Power aware task scheduling

A lot of work has developed DVFS for task scheduling. For example, Yao et al. [20] and Ali et al. [21] discuss scheduling independent tasks with DVFS on a single processor, Wei et al. [22] and Gruian et al. [23] use DVFS to schedule dependent tasks on multiple processors, Martin et al. [24] and Luo et al. [25,26] developed a power aware task scheduling algorithm for real time systems. As our work is devoted to developing power aware scheduling algorithms for dependent tasks, we compare our work with related research in this topic.

Zhang et al. [27], Martin et al. [24], Schmitz [28], and Luo et al. [26] schedule dependent tasks on real time, where the tasks