Accepted Manuscript

Dynamic energy-aware scheduling for parallel task-based application in cloud computing

Fredy Juarez, Jorge Ejarque, Rosa M. Badia

PII: S0167-739X(16)30214-X
DOI: http://dx.doi.org/10.1016/j.future.2016.06.029
Reference: FUTURE 3095

To appear in: *Future Generation Computer Systems*

Received date: 1 December 2015
Revised date: 20 April 2016
Accepted date: 23 June 2016

Please cite this article as: F. Juarez, J. Ejarque, R. M.Badia, Dynamic energy-aware scheduling for parallel task-based application in cloud computing, *Future Generation Computer Systems* (2016), http://dx.doi.org/10.1016/j.future.2016.06.029

Dynamic Energy-Aware Scheduling for Parallel Task-Based Application in Cloud Computing

Fredy Juarez[a,b], Jorge Ejarque[a,*], Rosa M. Badia[a,c]

[a]Barcelona Supercomputing Center (BSC), Workflows and Distributed Computing Group , Barcelona, Spain, 08034
[b]Instituto Tecnológico Superior de Álamo Temapache, Xoyotitla, Veracruz, México 92730
[c]Artificial Intelligence Research Institute (IIIA), Spanish National Research Council (CSIC)

Abstract

Green Computing is a recent trend in computer science, which tries to reduce the energy consumption and carbon footprint produced by computers on distributed platforms such as clusters, grids, and clouds. Traditional scheduling solutions attempt to minimize processing times without taking into account the energetic cost. One of the methods for reducing energy consumption is providing scheduling policies in order to allocate tasks on specific resources that impact over the processing times and energy consumption. In this paper, we propose a real-time dynamic scheduling system to execute efficiently task-based applications on distributed computing platforms in order to minimize the energy consumption. Scheduling tasks on multiprocessors is a well know NP-hard problem and optimal solution of these problems is not feasible, we present a polynomial-time algorithm that combines a set of heuristic rules and a resource allocation technique in order to get good solutions on an affordable time scale. The proposed algorithm minimizes a multi-objective function which combines the energy-consumption and execution time according to the energy-performance importance factor provided by the resource provider or user, also taking into account sequence-dependent setup times between tasks, setup times and down times for virtual machines (VM) and energy profiles for different architectures. A prototype implementation of the scheduler has been tested with different kinds of DAG generated at random as well as on real task-based COMPSs applications. We have tested different size instances and importance factors, and have evaluated which combination provides a better solution and energy savings in each case. Moreover, we have also evaluated the introduced overhead by measuring the time for getting the scheduling solutions for a different number of tasks, kinds of DAG, and resources, concluding that our method is suitable for run-time scheduling.

Keywords: Distributed Computing, Cloud Computing, Green Computing, Task-based Applications, Energy-aware Scheduling, Multi-Heuristic Resource Allocation, Makespan, Total Energy Flow

1. Introduction

Recent studies [1] [2] have estimated that around 1.5-2.0% of the total energy consumption is consumed by data centers, and this energy demand is growing extremely fast due to the popularization of Internet services and distributed computing platforms such as clusters, grids, and clouds. Regarding the efficiency of data centers, studies have concluded that, in average, around 55% of the energy consumed in a data center is consumed by the computing system and the rest is consumed by the support system such as cooling, uninterrupted power supply, etc. For that reason, green cloud computing is essential for ensuring that the future growth of cloud computing is sustainable [3].

There are several ways to reduce the energy consumed by an application when executed on a distributed platform: It includes the usage of low-power processor architectures or dynamic voltage frequency scaling (DVFS) [4], re-design of algorithms using energy-efficient patterns in compilers [5] or changing the scheduling policies for task-based applications on the available resources [6]. Traditionally, scheduling techniques have tried to minimize the total execution time of an application(makespan - C_{max}) [7] without worrying about the energy consumed. However, there is a trade-off between energy consumed and the execution time, and sometimes increasing the performance for a

*Corresponding author
Email address: jorge.ejarque@bsc.es (Jorge Ejarque)