Budget-constrained Resource Provisioning and Scheduling Algorithms for Scientific Workflows in Cloud Environments

[Extended Abstract]

Muhammad Hafizhuddin Hilman, Maria Alejandra Rodriguez and Rajkumar Buyya
CLOUDS Laboratory, School of Computing & Information System, The University of Melbourne
hilmanm@student.unimelb.edu.au, {marodriguez, rbuyya}@unimelb.edu.au

ABSTRACT

The deployment of large scale scientific workflow into cloud computing is popularly explored. Infrastructure as a Service (IaaS) Clouds offer flexible access to the resources (i.e., virtual machines) required for the execution of scientific workflows. This flexibility leads to a trade-off between two conflicting Quality of Service (QoS) requirements: time and cost. Majority of works in this area focus on meeting the deadline while minimizing the cost and few of them exploiting the budget spending to minimize the makespan. In this paper, we propose the Fastest-First Task-based Distribution (FFTD) and Slowest-First Task-based Distribution (SFTD), budget distribution algorithms designed to minimize the makespan with budget constraint. Our simulation results demonstrate that FFTD can meet the budget constraint and achieve lower makespan when compared to state-of-the-art budget distribution algorithm.

CCS Concepts

•Networks → Cloud computing; •Software and its engineering → Scheduling;

Keywords

Scientific Workflows; Resource Provisioning; Scheduling; Cloud Computing

1. INTRODUCTION

E-Science experiments have been drawing people's attention after The LIGO project successfully found gravitational waves. These experiments are generally composed of multiple scientific applications linked together, as workflows. They are large-scale applications and cloud computing has become popular for their deployment in particular. However, this leads to a trade-off between two important QoS parameters: time and cost. Thus, provisioning algorithms that can decide the type and number of VMs required by a workflow execution and scheduling algorithms capable of efficiently mapping the tasks to the resources while considering time and cost are essential. We propose the Fastest-First Task-Based Distribution (FFTD) and Slowest-First Task-Based Distribution (SFTD), budget distribution algorithms that aim to meet the budget constraint and minimize the makespan by distributing a portion of the budget to individual tasks and spending it only when idle time slots on VMs cannot be reused.

2. THE ALGORITHMS

We modify the EPSM algorithm [1] to a budget-constrained scheduling algorithm. Our budget distribution algorithms work by distributing the workflow budget into each individual tasks based on the execution order estimation. The order is defined by allocating the tasks into levels and sort them in each level based on The Earliest Finish Time (EFT). FFTD allocates the fastest resources within the budget to the earliest tasks while SFTD spends the slowest resources to as much tasks as possible and upgrade the allocation whenever all tasks got a share and still exists a leftover. The scheduling algorithm updates the status of budget spending whenever a task finished by recalculating the real cost and upgrading or downgrading the allocated budget for the following tasks.

3. RESULTS AND CONCLUSION

We compare our algorithms with Budget Distribution Trickling All In (BDT-AI) [2]. BDT-AI is a dynamic level-based algorithm that has the same objectives as our algorithms. We use CloudSim to model the cloud environment and a synthetic workflow with 100 tasks based on a real application from earthquake sciences (CyberShake) as a benchmark. To evaluate the performance, we use 10 different budget ranging from tighest to more relaxed ones. We use cost / budget ratio and makespan as a performance metric to evaluate the algorithms.

Generally, all algorithms are able to meet the budget constraint despite in the tighest budget, FFTD and BDT-AI violate the limit by an insignificant margin. In 80% of the cases, FFTD produced lower makespan than SFTD and BDT-AI. Meanwhile in the tighest budget, SFTD shown the lowest cost/budget ratio and makespan.

4. REFERENCES

[1] M. A. Rodriguez and R. Buyya, "Scheduling dynamic workloads in multi-tenant scientific workflow as a service platforms," *Future Generation Computer Systems*, 2017.

[2] V. Arabnejad, K. Bubendorfer, and B. Ng, "Budget distribution strategies for scientific workflow scheduling in commercial clouds," in *Proceedings of IEEE 12th International Conference on e-Science*. IEEE, 2016, pp. 137–146.