Accepted Manuscript

Scheduling deadline constrained scientific workflows on dynamically provisioned cloud resources

Vahid Arabnejad, Kris Bubendorfer, Bryan Ng

PII: S0167-739X(17)30020-1
DOI: http://dx.doi.org/10.1016/j.future.2017.01.002
Reference: FUTURE 3284

To appear in: *Future Generation Computer Systems*

Received date: 13 May 2016
Revised date: 10 October 2016
Accepted date: 7 January 2017

Please cite this article as: V. Arabnejad, K. Bubendorfer, B. Ng, Scheduling deadline constrained scientific workflows on dynamically provisioned cloud resources, *Future Generation Computer Systems* (2017), http://dx.doi.org/10.1016/j.future.2017.01.002

Scheduling Deadline Constrained Scientific Workflows on Dynamically Provisioned Cloud Resources

Vahid Arabnejad, Kris Bubendorfer, Bryan Ng

School of Engineering and Computer Science, Victoria University of Wellington, New Zealand

Abstract

Commercial cloud computing resources are rapidly becoming the target platform on which to perform scientific computation, due to the massive leverage possible and elastic pay-as-you-go pricing model. The cloud allows researchers and institutions to only provision compute when required, and to scale seamlessly as needed. The cloud computing paradigm therefore presents a low capital, low barrier to operating dedicated HPC eScience infrastructure. However, there are still significant technical hurdles associated with obtaining sufficient execution performance while limiting the financial cost, in particular, a naive scheduling algorithm may increase the cost of computation to the point that using cloud resources is no longer a viable option.

The work in this article concentrates on the problem of scheduling deadline constrained scientific workloads on dynamically provisioned cloud resources, while reducing the cost of computation. Specifically we present two algorithms, Proportional Deadline Constrained (PDC) and Deadline Constrained Critical Path (DCCP) that address the workflow scheduling problem on such dynamically provisioned cloud resources. These algorithms are additionally extended to refine their operation in task prioritization and backfilling respectively. The results in this article indicate that both PDC and DCCP algorithms achieve higher cost efficiencies and success rates when compared to existing algorithms.

Keywords: Scientific Workflows, Scheduling, Deadline Constrained, Cloud Resources

Email addresses: Vahid.Arabnejad@ecs.vuw.ac.nz (Vahid Arabnejad), kris@ecs.vuw.ac.nz (Kris Bubendorfer), bryan.ng@ecs.vuw.ac.nz (Bryan Ng)