Future Generation Computer Systems 52 (2015) 1–12

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Taxonomies of workflow scheduling problem and techniques in the cloud

Sucha Smanchat *, Kanchana Viriyapant

Faculty of Information Technology, King Mongkut's University of Technology North Bangkok, 1518 Pracharat 1 Road, Bangsue, Bangkok, Thailand

HIGHLIGHTS

- We propose taxonomies of cloud workflow scheduling problem and techniques.
- Aspects and classifications unique to cloud workflow scheduling are identified.
- Several techniques are reviewed and classified based on the proposed taxonomies.
- Some issues of future concern in cloud workflow scheduling are discussed.

ARTICLE INFO

Article history:
Received 12 January 2015
Received in revised form
25 March 2015
Accepted 30 April 2015
Available online 8 May 2015

Keywords:
Cloud computing
Cloud workflow
Workflow scheduling
Workflow scheduling taxonomy

ABSTRACT

Scientific workflows, like other applications, benefit from the cloud computing, which offers access to virtually unlimited resources provisioned elastically on demand. In order to efficiently execute a workflow in the cloud, scheduling is required to address many new aspects introduced by cloud resource provisioning. In the last few years, many techniques have been proposed to tackle different cloud environments enabled by the flexible nature of the cloud, leading to the techniques of different designs. In this paper, taxonomies of cloud workflow scheduling problem and techniques are proposed based on analytical review. We identify and explain the aspects and classifications unique to workflow scheduling in the cloud environment in three categories, namely, scheduling process, task and resource. Lastly, review of several scheduling techniques are included and classified onto the proposed taxonomies. We hope that our taxonomies serve as a stepping stone for those entering this research area and for further development of scheduling technique.

1. Introduction

Scientific workflows have been employed to streamline computational experiments to be executed automatically so that scientists are relieved of technical impediment. To cater for the computational power required to execute large scale experiments, these workflows have successfully been executed in the grid computing environment, which is a distributed system based on resource sharing and coordination. Such execution is usually planned and scheduled so that the workflow is efficiently executed on grid resources. For this purpose, scheduling techniques for scientific grid workflows have been studied and developed over many years [1,2].

With the emergence of cloud computing, which offers computing resources elastically on demand [3], scientific workflows, like other computer applications, can benefit from virtually unlimited resources with minimal hardware investment. With such advantages, cloud computing has attracted much attention and the research in workflow scheduling has thus shifted to workflow execution in the cloud environment. Unlike grid resources, which are usually utilized free of charge according to sharing agreements, cloud resources such as virtual machines are associated with costs according to their usage. Also, resources in the cloud are provisioned in different ways. Therefore, workflow execution in the cloud needs a scheduling technique that is different from those used by the grid workflows.

Since the birth of cloud computing, several cloud workflow scheduling techniques have been proposed. These techniques consider many aspects of cloud environment, which can be different due to the flexibility of cloud computing. For example, some techniques assume that a workflow is executed in hybrid cloud environment while others assume only a public cloud platform. These

* Corresponding author.
E-mail addresses: sucha.smanchat@acm.org, sucha.s@it.kmutnb.ac.th (S. Smanchat), kanchana.v@it.kmutnb.ac.th (K. Viriyapant).

http://dx.doi.org/10.1016/j.future.2015.04.019
0167-739X/

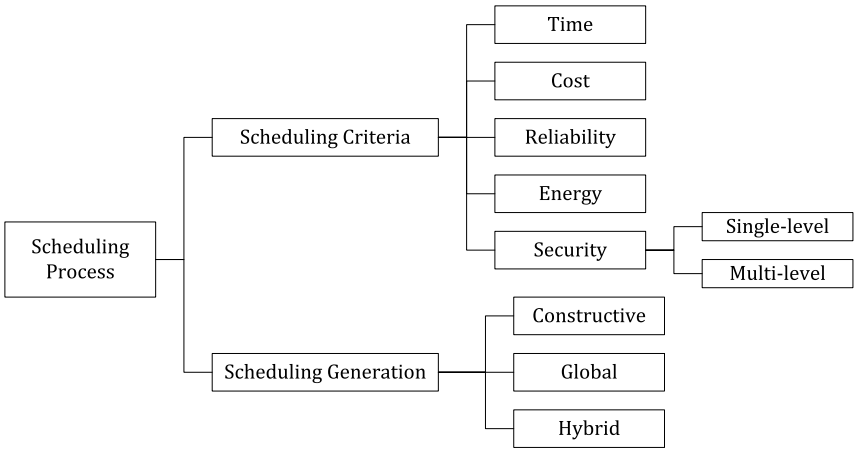

Fig. 1. Taxonomy of cloud workflow scheduling process.

different considerations influence the workflow scheduling problem leading to the techniques of different designs.

This paper aims to identify these aspects of consideration in cloud workflow scheduling in the form of taxonomies. In the past, Wieczorek, Hoheisel, and Prodan [4,5] presented taxonomies of workflow scheduling problem in the grid computing domain. Although, many aspects and classifications are shared in both the grid and the cloud contexts, new considerations arise in scheduling cloud workflow. Bittencourt, Madeira, and Fonseca [6] identified some issues of concern when scheduling in cloud environment but they focus mainly on resource characteristics and hybrid cloud execution. Alkhanak, Lee, and Khan [7] proposed a classification of challenges in cloud workflow scheduling focusing more on scheduling objectives and functionalities of workflow system architecture.

Other literature surveys of cloud workflow scheduling mostly focus only on the scheduling objective and the nature of each technique, which do not provide a comprehensive view of the research in this field [8–11]. In this paper, we propose taxonomies that complement the taxonomies of grid workflow scheduling in [4] to accommodate new issues introduced by cloud computing. In our review of cloud workflow scheduling literature, several new aspects can be identified and categorized, such as different types of resource usage costs and provisioning.

The structure of this paper is as follows. Section 2 defines a generic definition of cloud workflow scheduling problem. Section 3 defines and explains the aspects and the classifications that are used to form our taxonomies of cloud workflow scheduling problem and techniques. Some notable techniques are presented in Section 4 along with a comparison summary. The paper is then concluded in Section 5.

2. Cloud workflow scheduling problem

This section presents generic definitions related to cloud workflow scheduling; different techniques may use different forms of these definitions. Generally, a workflow is modeled as a directed acyclic graph (DAG). A workflow is defined as $w = (T, E)$ where T is a set of tasks represented by vertices in the DAG and E is a set of precedent dependencies represented by edges in the DAG. A task dependency is defined as $e = (t_i, t_j)$ where $t_i \in T$ is a predecessor task of $t_j \in T$ and $t_i \neq t_j$. A task t can start its execution if and only if all of its predecessor tasks have completed their execution.

Cloud workflow scheduling is then the mapping $T \rightarrow R$ from each task $t \in T$ to a resource $r \in R$ so that the specified criteria are met. In this domain, resources in the cloud are usually assumed to be virtual machine instances that are instantiated from machine images readily stored in the cloud (i.e. the concerns related to moving the machine images to the cloud is excluded). A task may require a program or software to execute, which are assumed to be preinstalled on the machine images. Therefore, it is also assumed that a task can be executed by any virtual machine instance type (i.e. any machine specification) by instantiating from the corresponding machine image [12].

Thus, a task t can be defined as $t = (id, MI)$, where id is the identity of t, and $mi \in MI$ is a set of virtual machine images that can perform execution for t. A compute resource r can be defined as $r = (mi, mt)$, where $mt \in MT$ is the virtual machine instance type of r among the instance types in MT that are offered by cloud providers.

3. Taxonomies of cloud workflow scheduling problem and techniques

This section identifies and explains the aspects and classifications of workflow scheduling in the cloud environment. Some of these aspects were identified in the taxonomies of grid workflow scheduling presented by Wieczorek, Hoheisel, and Prodan [4] and in a review by Bittencourt, Madeira, and Fonseca [6]. Due to the different nature of the grid and of the cloud in resource provisioning, new considerations in the scheduling problem and techniques arise. Also, as an initial work, we have identified a set of information required to implement a cloud workflow scheduler in [13]. To avoid reinventing the wheel, our taxonomies complement and, where possible, augment the taxonomies of grid workflow scheduling in [4].

The aspects presented herein are extracted mainly from research literature in cloud workflow scheduling domain and are separated into three main groups namely: *scheduling process*, *task* and *resource* partially following the taxonomies in [4]. Since cloud computing introduces the new compute resource paradigm, the concentration is more on the aspects in the *resource* group than the other two.

3.1. Taxonomy of scheduling process

This group contains aspects related to the scheduling process of scheduling algorithms, which extend the taxonomy in [4]. These aspects are shown in Fig. 1. Apart from our classification, Malawski et al. [11] also differentiates whether virtual machine instances are provisioned statically or dynamically during the scheduling process.