Future Generation Computer Systems 50 (2015) 3–21

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Cost-aware challenges for workflow scheduling approaches in cloud computing environments: Taxonomy and opportunities

Ehab Nabiel Alkhanak *, Sai Peck Lee, Saif Ur Rehman Khan

Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia

HIGHLIGHTS

- Extensively reviews cost-aware workflow scheduling approaches in cloud computing.
- Presents a set of taxonomy for cost-aware workflow scheduling challenges.
- Critically analyzes reported cost-aware workflow scheduling challenges.
- Provides useful recommendations for service consumers and utility providers.

ARTICLE INFO

Article history:
Received 25 May 2014
Received in revised form
10 December 2014
Accepted 19 January 2015
Available online 2 February 2015

Keywords:
Workflow scheduling
Cost benefit analysis
Quality of service
System architecture
Cloud computing
Taxonomy

ABSTRACT

Workflow Scheduling (WFS) mainly focuses on task allocation to achieve the desired workload balancing by pursuing optimal utilization of available resources. At the same time, relevant performance criteria and system distribution structure must be considered to solve specific WFS problems in cloud computing by providing different services to cloud users on pay-as-you-go and on-demand basis. In the literature, various WFS challenges affecting WFS execution cost have been discussed. However, prior work did not consider such challenges collectively. The main objective of this paper is to facilitate researchers in selecting appropriate cost-aware WFS approaches from the available pool of alternatives. To achieve this objective, we conducted an extensive review to investigate and analyze the underlying concepts of the relevant approaches. The cost-aware relevant challenges of WFS in cloud computing are classified based on Quality of Service (QoS) performance, system functionality and system architecture, which ultimately result in a taxonomy set. Some research opportunities are also discussed that help in identifying future research directions in the area of cloud computing. The findings of this review provide a roadmap for developing cost-aware models, which will motivate researchers to propose better cost-aware approaches for service consumers and/or utility providers in cloud computing.

1. Introduction

Workflow Scheduling (WFS) is aimed to automatically assign and manage the execution of dependent tasks on shared resources that are controlled by a workflow scheduler [1,2]. Researchers focus on WFS due to its ability to handle complicated applications (e.g. drug discovery and weather modeling [3]) with the whole or part of complex task dependencies.

WFS in cloud computing allocates each task to a relevant cloud service by ordering the execution of various resources to obtain satisfactory Quality of Service (QoS). Researchers have mentioned three main benefits of migrating workflow to cloud computing [4–12]: (i) enable the utilization of various cloud services to facilitate the automation of distributed large-scale workflow execution; (ii) significant reduction of hardware expenditure for workflow execution by sharing and providing resources in cloud systems; and (iii) increased user satisfaction along with reduced execution cost and time by obtaining the "pay-as-you-go" business model. Furthermore, in a cloud computing environment, the Workflow Management System (WFMS) belongs to a specific service provider and is under the management of relevant cloud resource manager(s). Thus, WFMS provides the ability to get access to other cloud services and facilitates the Service Level Agreements (SLA) [13,14].

One of the most challenging endeavors of WFMS is to provide different services to cloud users pay-as-you-go and on-demand basis. Consequently, this endeavor introduces a number of problems: (i) different users compete for resources within the cloud computing environment [15–17]; (ii) data needs to be transferred from one

* Corresponding author.
E-mail addresses: ehabsoa@siswa.um.edu.my (E.N. Alkhanak), saipeck@um.edu.my (S.P. Lee), saif_rehman@siswa.um.edu.my (S.U.R. Khan).

http://dx.doi.org/10.1016/j.future.2015.01.007
0167-739X/

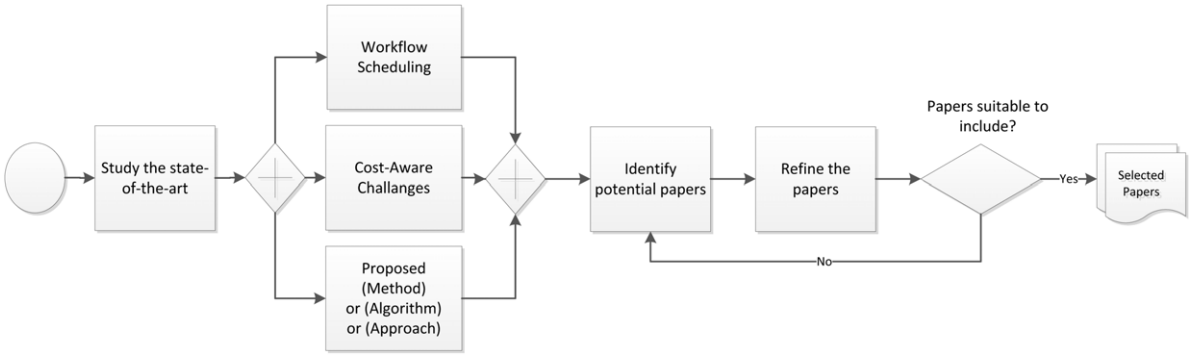

Fig. 1. Methodological approach of paper selection.

resource to another [18–23]; (iii) the inter-dependency between tasks introduces high communication cost and/or time [24–30]; and (iv) the trade-off between QoS of user requirements and the cost of workflow execution [1,10,13,31–34]. However, this review targets execution cost-related problems which can possibly affect both service consumers and utility providers.

Generally, the WFS problem is known to be NP-complete [8,35]. Since no optimal solution can be found for the NP-complete problem, it has drawn researchers to propose a plethora of cost-aware approaches with the aim of providing a near-optimal solution [33,36]. However, current state-of-the-art studies tackle different scheduling problems in cloud workflow systems by focusing on general QoS optimization constraints or specific workflow applications [37–47]. To the best of our knowledge, none of these studies provide an integrated view of cost challenges. As a result, there is a research gap in dealing with interrelationships among the cost-aware WFS approaches, and QoS, system functionality and system architecture challenges.

The key objective of this paper is to identify the cost-aware challenges of WFS in cloud computing. To achieve this objective, we targeted to determine the following secondary objectives:

(a) to analyze state-of-the-art cost-aware WFS approaches based on three important aspects (i.e. QoS, system functionality, and system architecture),

(b) to devise relevant taxonomies of cost-aware WFS challenges (i.e. cost-aware WFS challenges, cloud workflow system aspects, and cost-aware WFS profitability),

(c) to find the correlation between the WFS cost-aware challenges and profitability, and

(d) to identify the future opportunities in this field of research that helps in suggesting new research directions in WFS.

To achieve the mentioned objectives, we formulated the following research questions:

RQ1: What are the key cost-aware challenges of WFS?

Answering RQ1 would help researchers understand the reported cost-aware challenges in grid and cloud computing. Thus, it would help with devising taxonomies, which would consequently enable researchers to achieve minimum execution cost with WFS in cloud computing (Section 2).

RQ2: What are the state-of-the-art approaches of cost-aware WFS?

Answering RQ2 would provide researchers with a description of underlying techniques, features, computing environment, and type of approaches for each of the reported approaches (Section 3).

RQ3: How the challenges related to QoS, system functionality, and system architecture are affected by the execution cost of WFS in cloud computing?

Answering RQ3 would provide researchers with a complete description and analysis of the cost-aware challenges in cloud workflow system from the QoS, system functionality, and system architecture (Section 4).

RQ4: How is the profitability of cost-aware WFS affected by QoS, system functionality, and system architecture challenges?

Answering RQ4 would help researchers identify the relevant cost-aware approaches based on the purpose of the model that might be affected by the service consumers and/or utility providers. Additionally, the results of applying taxonomies on the reviewed cost-aware approaches will be obtained through the complete relation between the challenges of QoS, system functionality and system architecture (Section 4).

1.1. Motivation

Cloud computing is an emerging area of research for the past few years. A significant number of approaches have been proposed to handle different WFS challenges. The cost challenge remains an important consideration due to its direct impact on both service consumers and utility providers from different aspects (i.e. QoS, system functionality, and system architecture).

Several reviews have been conducted that cover various types of WFS challenges in cloud and grid computing. Yu and Buyya [42] classified and characterized a number of approaches for structuring and executing workflows in grid computing. Later, Yu and Buyya [43] reviewed a number of grid computing workflow management systems. Next, Prodan and Wieczorek [6,37] devised a taxonomy of multi-criteria problem for WFS. Their devised taxonomy classified the multi-criteria into five different aspects based on the workflow structure (i.e. cost aggregation method, intra-dependence, optimization direction, and interdependence). Some work has emphasized on the WFS problem in cloud computing and grid computing [44,46,48]. Singh and Singh [40] reviewed diverse WFS algorithms and compared the algorithms according to their type, objective criteria, and environment. However, none of the published work has focused on the cost-aware WFS challenges. Furthermore, to the best of our knowledge, prior work lacks in providing a complete taxonomy of cost-aware WFS challenges and a thorough comparison of WFS approaches. Consequently, a comprehensive taxonomy is required to provide an in-depth understanding of cost-aware WFS challenges and opportunities that can be useful for future researchers.

1.2. Selection of papers

This section presents our methodological approach (Fig. 1) that we adopted to select the relevant papers. First of all, we devised a search query based on the formulated research questions:

[(workflow scheduling) AND ((cost-aware) AND (challenges)) AND ((cloud computing) OR (grid computing)) AND ((approaches) OR (models)) AND ((profitability) AND (service consumers) OR (utility providers))]

After that, we executed the search query on a number of search databases (IEEE, Elsevier, Springer, ACM, Google Scholar). Notice