Computer Networks 108 (2016) 184–198

Contents lists available at ScienceDirect

Computer Networks

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

Dynamic embedding of workflow requests for bandwidth efficiency in data centers☆

Tram Truong-Huu[a,*], Mohan Gurusamy[a], Vishal Girisagar[b]

[a] Department of Electrical and Computer Engineering, National University of Singapore, 4 Engineering Drive 3, Singapore 117583, Singapore
[b] Continental Automotive Singapore Pte Ltd, 80 Boon Keng Rd, Continental Building, Singapore 339780, Singapore

ARTICLE INFO

Article history:
Received 11 March 2016
Revised 15 July 2016
Accepted 18 August 2016
Available online 20 August 2016

Keywords:
Workflows
Workflow embedding
Cloud data centers
Bandwidth minimization
Critical path
WDM-based optical networks

ABSTRACT

In this paper, we address the problem of embedding dynamically-arriving workflow requests in data centers. Workflows pose challenges due to data precedence and time disjointness among tasks, thus driving the need for intelligent methods to embed workflows in data centers while improving the bandwidth efficiency as well as guaranteeing the application performance. We first formulate an integer programming optimization model for the embedding problem that minimizes the amount of bandwidth required for workflow execution. We then develop two algorithms namely Critical Path Workflow Embedding (CPWE) and Edge Priority Workflow Embedding (EPWE) to solve this problem. We consider two data center network architectures: packet switching electrical networks and circuit switching optical wavelength division multiplexed (WDM) networks. While WDM-based optical networks have much larger bandwidth capacity to meet the ever-growing traffic demand in data centers, they pose challenges due to wavelength continuity constraint and the nature of circuit switching. We thus additionally propose methods for selecting appropriate Top-of-the-Rack (ToR) switches and wavelengths during the embedding process so as to increase the chance of accommodating many requests that span over multiple ToRs. We evaluate CPWE and EPWE through comprehensive simulations. The results show that CPWE and EPWE significantly reduce the bandwidth required for a workflow request by up to 66% for random workflows and 80% for realistic-application workflows compared to baseline algorithms. The results also show that the proposed methods for ToR selection and wavelength selection in optical data centers outperform other methods by reducing the rejection ratio by up to 47% with dynamic reconfiguration of lightpaths and 40% with incremental configuration of lightpaths.

1. Introduction

Cloud data centers have become an attractive candidate to meet resource demands of individual users and institutions. Instead of setting up a local infrastructure that costs a lot for device purchases and maintenance services, users nowadays are exploiting public clouds, which provide immense computing capacity, quasi-unlimited storage space and broad access network connections. Many large-scale and data-intensive applications have been migrated to the cloud to be able to handle a big amount of data as well as heavy computations. These applications generate a huge traffic demands mainly remaining within data centers as reported in [2]. This huge traffic demands force data centers look for the means for not only increasing the network capacity but also improving network resource utilization. While electrical packet-switched networks still keep their important role, using WDM-based optical networks in data centers is becoming a trend to meet the network capacity challenge. In comparison with electrical packet-switched networks, WDM-based optical networks have much larger bandwidth capacity with low power consumption and cabling complexity. Dynamic reconfiguration of lightpaths also brings in the flexibility for network management and traffic engineering. However, in addition to the limitations of circuit-switching, the wavelength continuity constraint due to the dynamic arrival of resource requests and high degree of the optical switch, which interconnects multiple ToR switches, has high impact on resource usage efficiency, making the problem of request embedding more challenging. Thus, intelligent methods for ToR and wavelength selection are needed to embed resource requests so as to use the servers under ToRs and bandwidth (or wavelength) resources between ToRs efficiently.

☆ This article is an extended version of the paper published in ICDCN 2016 [1] that was carried out when Vishal Girisagar was affiliated with NUS.

* Corresponding author.
E-mail addresses: eletht@nus.edu.sg (T. Truong-Huu), elegm@nus.edu.sg (M. Gurusamy), vishalgirisagar801@gmail.com (V. Girisagar).

http://dx.doi.org/10.1016/j.comnet.2016.08.018
1389-1286/

Considering both electrical and optical network data centers, in this paper, we address the problem of embedding workflow requests, which usually require large amount of computing resources and guaranteed bandwidth for applications such as detecting gravitational-waves [3], weather forecasting [4], predicting earthquakes [5]. A workflow resource request is represented by a set of computing tasks and a set of edges that represent the data dependencies among the tasks. Each computing task requires a number of virtual machines (VMs) to run the application for a certain duration and each edge requires a specific network bandwidth for data transmission between the VMs allocated to dependant tasks. Unlike the traditional virtual network requests [6], in which all requested resources including bandwidth and virtual machines are needed during the entire lifetime of the request, workflow requests have specific resource requirements at different times. Due to the task dependencies, some workflow tasks can execute concurrently while others have to run sequentially. VMs and bandwidth requested for a task may not be utilized before its preceding tasks finish their execution. Furthermore, due to large resource demands, workflow tasks are likely to be embedded across multiple ToRs of data centers, making the impact of the wavelength continuity constraint more pronounced in optical data centers. Thus, embedding of workflow requests in optical data centers needs to consider not only the lifetime of the tasks but also the ToR and wavelength selection to ensure high connectivity between ToRs with wavelength-continuous paths.

Given a workflow resource request, cloud providers need to embed this request in physical servers in data centers, considering the capacity of physical servers and network links as well as the network topology. The embedding process consists of allocating VMs on physical servers under ToR switches and reserving bandwidth on the links of the data center to guarantee the performance of workflow applications. While the amount of computing resources cannot be optimized, i.e., the number of VMs requested needs to be always satisfied, the bandwidth consumption for workflow execution is affected by the embedding solution. If a physical server has sufficient VMs to host all the requested VMs of two adjacent tasks in the workflow, then the execution of these two tasks will not consume link bandwidth. Otherwise, a certain amount of bandwidth of the link connecting two physical servers hosting the VMs of the two tasks needs to be reserved. Since the network bandwidth is limited, minimizing the bandwidth consumption during the execution of a workflow allows providers to accept more workflow resource requests, thereby increasing the revenue from the users who pay for resource usage. An intelligent embedding technique is therefore needed to help providers optimize the resource utilization in data centers.

Since workflow requests are different from the traditional virtual network requests due to data precedence and time disjointness among workflow tasks, existing virtual network embedding algorithms such as those presented in [7,8] are no longer applicable for embedding workflow requests in data centers. In this paper, we present a novel model for embedding workflow resource requests in data centers, considering the above challenges. We first formulate an integer programming optimization problem, which aims at minimizing the amount of bandwidth required for the execution of a workflow while guaranteeing its computing and network resource demands. Solving such an optimization problem and its variations has been shown to be computationally hard due to their $\mathcal{NP}$-complete nature [9]. We therefore develop two heuristic algorithms namely Critical Path Workflow Embedding (CPWE) and Edge Priority Workflow Embedding (EPWE) to solve the embedding problem efficiently.

Algorithm CPWE tries to embed all the VMs requested by the tasks on the *critical path* of the workflow on to the same physical server. We define the critical path as the path from the entry task to the exit task of the workflow for which the total bandwidth consumed on the links is the highest. Algorithm CPWE therefore needs an initial phase that determines the critical path before doing the embedding. Algorithm EPWE is computationally simpler than CPWE. It first sorts the edges in the workflow in the descending order of bandwidth requirement, and then starts embedding the tasks following the specified order. In both cases, it may not be successful to embed the entire critical path or the edge with the highest bandwidth requirement on the same physical server. Algorithms CPWE and EPWE then try to embed the requests in different servers so as to reduce the bandwidth consumption. To address the specific challenges of optical networks, we propose a method for ToR selection based on a connectivity-index function, which defines how well a ToR is connected with other ToRs with free wavelength-continuous paths so as to increase the chance of accommodating many future requests that span over multiple ToRs. We also propose a function that computes the goodness value for each wavelength between a pair of ToRs to choose the best wavelength for a lightpath. These two methods are integrated into CPWE and EPWE when realizing workflow request embedding in optical data centers. We evaluate the performance of the proposed algorithms through comprehensive simulations and compare their performance against baseline algorithms to demonstrate their effectiveness.

The rest of the paper is organized as follows. We discuss the related works in Section 2. We present the system model and mathematical formulations in Section 3. We present the proposed algorithms in Section 4. We present the methods for ToR and wavelength selection in Section 5. We carry out performance study and analyze the simulation results in Section 6 before concluding the paper in Section 7.

2. Related work

2.1. Workflow execution in clouds

Recently, research on workflows has received significant attention specially for the complex applications, which require large amount of data and computationally complex resources. Various systems such as MOTEUR [10] and Kepler [11] are used to interpret the workflow applications and submit the workflow tasks to a computing infrastructure for execution. These systems request only the computing resources for the workflows but they do not consider the bandwidth requirements between the dependent tasks in a workflow. In [12], the authors focused on cost minimization for embedding workflow application on data centers but they also did not consider bandwidth requirements. In [13], the authors aimed at minimizing the amount of computing resources, i.e., the number of VMs needed to execute workflows while ensuring their execution deadline. Depending on the number of remaining tasks and the available time before the deadline, the algorithm dynamically scales the number of VMs: increasing the number of VMs to meet the deadline or decreasing the number of VMs to reduce the usage cost. This work assumed that the required bandwidth is available on the links. This may not be practical in data centers since multiple tenant requests arrive at the same time and share the residual bandwidth on the links.

2.2. Workflow embedding and scheduling

Workflow embedding and scheduling has received significant attention from research communities [12,14–17]. In [16], the authors presented the *Myopic* algorithm that maps each individual task of the workflow on a first suitable computing resource. This simple algorithm did not consider the dependencies between workflow tasks, leading to high bandwidth consumption. Three