Workflow-based Distributed Computing over Optical Virtual Private Networks

Yaohui Jin, Yan Wang, Wei Guo and Weisheng Hu
State Key Laboratory of Advanced Optical Communication Systems and Networks
Shanghai Jiao Tong University, 200240, China
Email: *jinyh@sjtu.edu.cn*
(Invited)

Abstract
It is beneficial, for both customers and carriers, to provide optical virtual private network for workflow-based distributed computing applications. We formulate the schedule system and also propose a rescheduling scheme under dynamic shared OVPN scenario.

1. Introduction

By using grid technologies, scientists expect to be able to share, select, and aggregate a wide variety of geographically distributed computational resources (e.g. supercomputers, data sources, storage systems, instruments etc) to solve scientific problems in physics, astronomy, biology, earthquake science, etc. These applications are often structured as workflows that express an application (or a job) by specifying a set of interdependent tasks. The inter task communication will generate or transfer large datasets frequently, so the availability of large amounts of bandwidth coupled with low latency is required. As the best-effort nature of traditional IP networks cannot supply sufficient QoS guarantees, optical circuit-switched (OCS) networking technologies are considered better suited to support such advanced scientific applications [1].

A hot research topic is optimal co-scheduling of computing resources and network resources, i.e., mapping tasks to computing facilities as well as lightpath establishment over optical network. This problem has been studied recently [2][3][4]. Most algorithms proposed in previous work are mainly based on static scheduling strategies with assumption that detail resource information and accurate performance prediction is available for both computation and communication costs. However, in many operational cases, network carriers may not want to expose too much detailed internal network information to customers due to confidentiality consideration.

In this paper, we employ Optical Virtual Private Network (OVPN), which is also referred to as layer one VPN (L1VPN) [5], to solve this problem. Both customers and carriers can benefit from such architecture. Carriers provide virtualized or partial network resource to customers, and customers can make co-scheduling of their own virtualized computing and networking resources with better isolation and scalability. In order to deal with the dynamic changes of both computing and OVPN resources, we propose a rescheduling mechanism which generates a near optimal schedule using a static scheduling algorithm for the job before it starts execution and then changes the initial schedule dynamically during the execution.

2. System Model and Problem Statement

A workflow application can be modeled as a directed acyclic graph (DAG) [2]. We formulate the task model as $G_t = (\mathbf{V}, \mathbf{E})$, where $\mathbf{V}$ is a set of v tasks and $\mathbf{E}$ is a set of e edges between the tasks. Each edge $e_{mn} \in \mathbf{E}$ represents the precedence constraint such that task v_n can not start execution until v_m finishes. The weight $w(e)$ denotes the data volume transmitted on the edge e. An example DAG is shown in Fig.1.
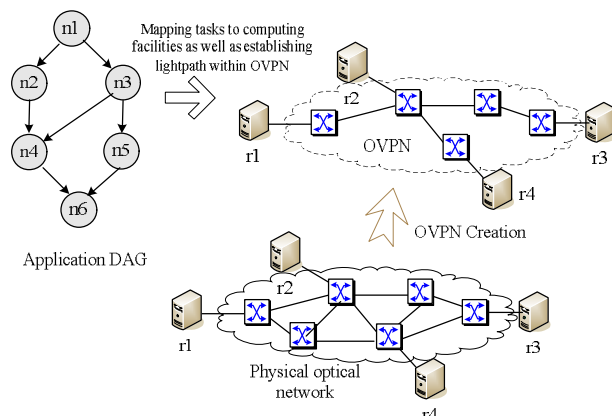

Fig. 1. Scheduling system under OVPN scenario

Our target problem can be stated as communication aware DAG scheduling problem, which consists of two sub problems under OVPN scenario (as shown in Fig. 1): 1) OVPN creation or design according selected computing resource group; 2) co-scheduling of task allocation and lightpath establishment within OVPN with the objective to minimize the schedule length (the execution time of DAG).

Since it is almost impossible to predict the traffic characteristics generated by the workflow based applications, we employ hose model of traffic demand and tree structure to design OVPN topology [6].

The OVPN resource management can be dedicated or shared. When the computing resource and the OVPN resource are both dedicated, we can directly employ the previous static joint scheduling method. If the computing resource or the OVPN run in a dynamic shared manner, in order to deal with the resource performance variation, we employ rescheduling mechanism which generates a

near optimal schedule using a static scheduling algorithm for the job before it starts execution and then changes the initial schedule dynamically during the execution.

3. Rescheduling under Dynamic Shared OVPN

We propose a computation and communication delay aware rescheduling (C^2DAR) scheme, which improves the initial static schedule of a DAG by considering only selective tasks or lightpaths for rescheduling based on measurable properties. If the delay of current task or lightpath will impact the complete job finish time, then the reschedule is triggered.

In order to facilitate determination of the delay property for rescheduling decision, we first introduce a schedule result graph (SRG) concept. Based on the SRG, we can easily compute the maximum delay that can be tolerated in the execution time of a scheduled task or lightpath without affecting the overall schedule length.

The SRG definition and the C^2DAR algorithm are not shown due to paper size limitation.

4. Simulation Study

We map a randomly generated DAG onto 4 computing resources connected by a 16-node NSFNET optical network with background traffics, and compare the performance of communication aware DAG scheduling under three optical network provision scenarios: full optical network (ON), dedicated OVPN (OVPN-D) and shared OVPN (OVPN-S). All the results are averaged over 100 simulations.

Fig. 2 shows the blocking probability of scheduling over different network provision models under varying background traffic load. And Fig. 3 compares the scheduling results under different scenarios. As can be seen, although OVPN-D can produce shortest schedule length, it has two problems: it occupies most network resources and is not applicable when the background traffic load is higher than 300 (the block ratio is more than 90%). The ON model achieves less network resource usage with lowest blocking probability. However, it has difficulty in implementation as discussed before. By comparison, OVPN-S achieves relatively better performance with lower blocking ratio for low and median traffic load.
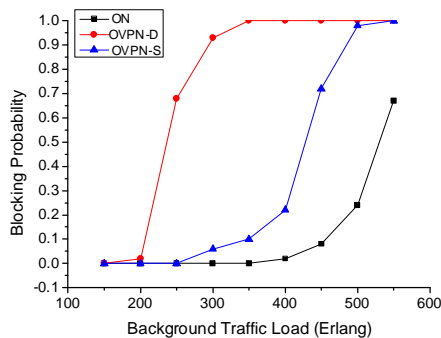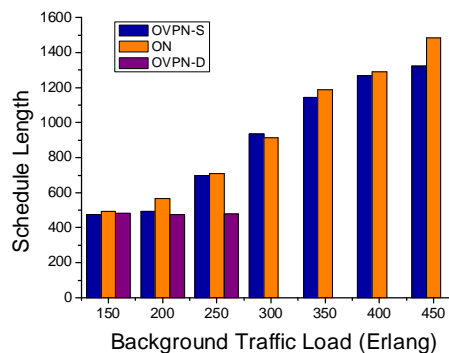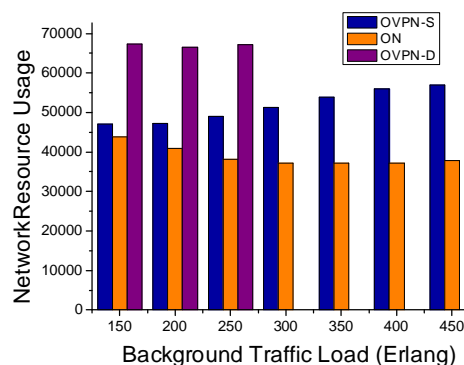
Fig. 2. Blocking probability of scheduling under ON, OVPN-D and OVPN-S scenarios respectively.

(a)

(b)

Fig.3. Comparison of (a) schedule length and (b) network resource usage with respect to background traffic load under 3 different network provision scenarios.

5. Conclusions

In this paper, we study why and how to employ OVPN for communication aware DAG scheduling problem. In order to deal with the dynamic changes of both computing and OVPN resources, we propose a C^2DAR rescheduling mechanism. Simulation results show that rescheduling under OVPN scenario achieves better performance.

References
1. D. Simeonidou, et al, IEEE Journal of Lightwave Technology, vol. 23, pp. 3347-3357, 2005.
2. Y. Wang, et al, Journal of Optical Networking, vol. 6, pp. 304-318, 2007.
3. Z. Sun, et al, IEEE J. Lightwave Technol. vol. 26, no. 17, pp.3011-3020, 2008.
4. X. Liu, et al, INFOCOM 2008. The 27th Conference on Computer Communications. IEEE , vol., no., pp.1966-1974, 13-18 April 2008
5. T. Takeda, et al, IETF RFC 4847, April 2007.
6. A. Kumar,et al, IEEE Trans. Networking, vol.10, no.4, pp.565-578, Aug. 2002.