J. Vis. Commun. Image R. xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Online multi-objective optimization for live video forwarding across video data centers ☆

Wu Liu [a,*], Yihong Gao [a], Huadong Ma [a], Shui Yu [b], Jie Nie [c]

[a] Beijing Key Lab of Intelligent Telecomm. Software and Multimedia, Beijing University of Posts and Telecomm., Beijing 100876, China
[b] School of Information Technology, Deakin University, Melbourne, VIC 3125, Australia
[c] Department of Computer Science and Technology, Tsinghua University, Beijing, China

ARTICLE INFO

Article history:
Received 26 September 2016
Revised 31 December 2016
Accepted 7 January 2017
Available online xxxx

Keywords:
Big video data
Video data center
Live video forwarding
Resource provisioning optimization

ABSTRACT

The proliferation of video surveillance has led to surveillance video forwarding services becoming a basic server in video data centers. End users in diverse locations require live video streams from the IP cameras through the inter-connected video data centers. Consequently, the resource scheduler, which is set up to assign the resources of the video data centers to each arriving end user, is in urgent need of achieving the global optimal resource cost and forwarding delay. In this paper, we propose a multi-objective resource provisioning (*MORP*) approach to minimize the resource provisioning cost during live video forwarding. Different from existed works, the MORP optimizes the resource provisioning cost from both the resource cost and forwarding delay. Moreover, as an approximate optimal approach, MORP adaptively assigns the proper media servers among video data centers, and connects these media servers together through network connections to provide system scalability and connectivity. Finally, we prove that the computational complexity of our online approach is only $O(\log(|U|))$ ($|U|$ is the number of arrival end users). The comprehensive evaluations show that our approach not only significantly reduces the resource provisioning cost, but also has a considerably shorter computational delay compared to the benchmark approaches.

1. Introduction

Recently, massive IP cameras in diverse locations are connected to the geo-distributed video data centers (VDCs) to provide live video streams for end users [1–3]. The surveillance video service [4–8] provided by the VDCs can be viewed as *Video Surveillance as a Service* (*VSaaS*). For example, travelers can remotely obtain views around the scenic locations by accessing IP cameras to plan their tours.[1] Furthermore, *VSaaS* can also extract the objects of interest from the video streams to help travelers recognize landmarks and special events on their trips [9–13]. In this service, travelers located at different places randomly access IP cameras deployed at several hot spots. The service requires both server and network resources of the video data centers to forward the live video streams. Thus, an online and cost-saving resource provisioning plan is the key point of the VDCs to serve the end users.

Generally, VDCs at different locations are connected by high-capacity networks [14]. There is typically more than one VDC around the arriving end user. As shown in Fig. 1, each VDC employs virtual machines (VMs) as media servers, which can provide the end users with the live video streams. A scheduler is set up to decide the media server and network resources of the VDCs for each end user. Furthermore, VDCs rent the media servers and bandwidths at different prices,[2] and employ different forwarding delays to forward the live video stream for each end user [23]. Consequently, the resource provisioning approach of the scheduler must not only reduce the resource provisioning cost but also consider the media server cost, the bandwidth cost, and the forwarding delay as the factors of the resource provisioning cost during resource provisioning, which is a multi-objective resource provisioning approach. Furthermore, since the requests arrive in real time, the approach must be an online approach, which can only obtain the local information of the resource utilization to provide resources for each arriv-

☆ This paper has been recommended for acceptance by Anan Liu.

* Corresponding author.
E-mail addresses: liuwu@bupt.edu.cn (W. Liu), mhd@bupt.edu.cn (Y. Gao), hii_gao@hotmail.com (H. Ma), shui.yu@deakin.edu.au (S. Yu), niejie@tsinghua.edu.cn (J. Nie).

[1] It is a service of MegaEyes. MegaEyes is a video surveillance system on a video data center. http://qqy.fjii.com/, Aug. 2016.

[2] For example, Amazon EC2, Aliyun, and Google Cloud Platform charge different resource prices for the video data centers from different locations.

http://dx.doi.org/10.1016/j.jvcir.2017.01.010
1047-3203/

Please cite this article in press as: W. Liu et al., Online multi-objective optimization for live video forwarding across video data centers, J. Vis. Commun. (2017), http://dx.doi.org/10.1016/j.jvcir.2017.01.010

ARTICLE IN PRESS

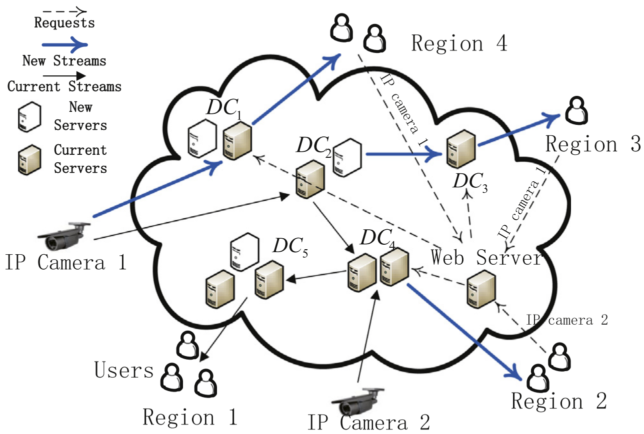

Fig. 1. System overview for the live video forwarding across VDCs.

ing end user. Therefore, the main challenge in providing resources for video stream forwarding is to achieve the global optimal resource provisioning cost with local information.

Existing video data centers provide three different services: video delivery service, interactive video forwarding service, and live video forwarding service. These three services provide different resource provisioning solutions. (1) For the video delivery service, most resource provisioning approaches utilize the P2P or server-client structure to manage the resources [19–21]. Nonetheless, the approaches are offline, which require the arrival sequence of the requests before resource provisioning. (2) For the interactive video forwarding service, the resource provisioning approaches provide media servers between any two communicating users [22–27]. However, the resource provisioning does not share the media servers for different communicating pairs, and thus, the resource provisioning approaches are very costly in terms of the interactive forwarding service. (3) For the live video forwarding service, the media servers are constructed as a forwarding tree from the end users to the video sources, and the forwarding tree can share the resources with different end users [28–32]. However, the resource scalability, prices and forwarding delay factors are not comprehensively considered during the resource provisioning for the forwarding tree. Finally, the online resource provisioning approaches across VDCs primarily maximize the throughput of the system [19,30]. The number of media servers in these approaches cannot be changed, and the arrival requests can be denied. In brief, the previous works cannot comprehensively optimize the resource cost and forwarding delay for online video forwarding.

In this paper, we propose an online resource provisioning approach, which is called multi-objective resource provisioning (*MORP*), to minimize the resource provisioning cost for *VSaaS*. Our approach views the forwarding tree as a route for the live video stream from IP cameras to all end users simultaneously. Different from the previous work in [28–32], we take the media server capacity, the resource scalability, the resource prices, and the forwarding delay into account. Since the end users randomly arrive and leave the media servers, our approach adaptively provides or removes media servers of the VDCs, and dynamically adjusts the forwarding trees for the end users. First, we formulate the resource provisioning problem as a route selection problem, which organizes the media servers at runtime to optimize the resource provisioning cost. The traditional mathematical methods for solving the route selection problem have a computational delay that is too long to assign the media servers and the network connections. To improve the computational delay, we convert the problem into two traditional resource provisioning problems with consideration of the resource capacity constraint, the resource prices, and the forwarding delay: *facility location* (FL) problem and *shortest path* (SP) problem. This conversion can obtain the approximate optimal solutions with a lower competitive ratio. For the *FL* problem, we design an online iterative algorithm to provide resource scalability, which chooses a proper media server among the VDCs. For the *SP* problem, we further design an online algorithm to support the connectivity with the minimum cost, which establishes a forwarding path from the chosen media server to the forwarding tree. We prove that our approach is $O(\log(|U|))$-competitive ($|U|$ is the number of end users), and the time complexity of our approach is $O(|F|\log(|F|))$ ($|F|$ is the number of VDCs). Finally, we provide a resource releasing approach to further save the resource cost. The experimental results show that the resource provisioning cost of VDCs is clearly reduced by our approach, and the computational delay of the approach is greatly reduced. Our work includes the following contributions:

- We propose an online multi-objective optimization approach for live video forwarding in video data centers. The proposed approach comprehensively considers the forwarding delay, server and bandwidth costs, and the availability and scalability of VDCs.
- Compared to the traditional approaches, the proposed approach achieves better time complexity ($O(|F|\log(|F|))$), and is feasible with an $O(\log(|U|))$ competitive ratio.
- The extensive experiments demonstrate the effectiveness and efficiency of our approach based on a real-world scenario.

The remainder of this paper is organized as follows. The system framework for live video forwarding is introduced in Section 2. In Section 3, we discuss the system models and the resource provisioning problem. Section 4 presents how to provide the resources for each arriving end user to minimize the resource provisioning cost. We further show the experimental results in Section 5. Section 6 provides the related work. Finally, the conclusions are given in Section 7.

2. System overview

There are four participants in Fig. 1 to provide the *VSaaS*: end users, Web server, VDCs, and IP cameras. When the end users require the live video streams from the IP cameras, the VDCs must assign a media server to forward the streams. The VDC can provide current or new media servers to serve the end users (for example, the DC_1 provides a new media server for the end user in region 4. The DC_4 utilizes the current media server to serve the end user from region 2). For the network usage, the serving media server can directly receive the required live video stream from the IP camera (for example, DC_1 and DC_4 build up connections to IP cameras to serve end users). The serving media server can also receive the live video streams from the media server of the neighboring VDCs (the relationship between DC_3 and DC_2). During these processes, some new media servers and network connections are employed. The details of each component are given as follows:

2.1. End users

Each end user accesses the Web server to require the live video stream of a certain IP camera and waits for the addresses of the chosen media servers. When the addresses are received, the end users connect to the media servers and obtain the live video streams of the required IP cameras.

Please cite this article in press as: W. Liu et al., Online multi-objective optimization for live video forwarding across video data centers, J. Vis. Commun. (2017), http://dx.doi.org/10.1016/j.jvcir.2017.01.010