Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
August 2014, 21(4): 47–53
www.sciencedirect.com/science/journal/10058885
The Journal of China Universities of Posts and Telecommunications
http://jcupt.xsw.bupt.cn

Cloud download system optimizing by job and notification scheduling

XU Ying-ying, CHEN Chang-jia, ZHAO Yong-xiang, CHEN Yi-shuai

School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing 100044, China

Abstract

With expansion of cloud-based applications, it is necessarily for business service provider to deal with how to provision resources to server customers. In cloud download, the service provider faces great challenge on determining when to process job and when to notify the user to meet the conflicting objectives, i.e., the service profit maximization and user satisfaction maximization. An economic model was proposed to find job and notification scheduling strategies for making tradeoffs between user satisfaction and profit of the system. In this model, optimization problems are formulated as follows, profit maximization under a satisfaction target and then satisfaction optimization with a profit bound. By solving these optimization problems, the optimal scheduling strategies are obtained under various quality requirements of service. It is demonstrated that the proposed strategy can obtain better performance than the system without the intelligent scheduling strategy.

Keywords cloud download, economic model, optimization, scheduling strategy, user satisfaction

1 Introduction

With cloud computing enable rapid provisioning server utilities to users anywhere, anytime, a growing number of applications is created on the cloud platform to provide cost-effective service, like cloud media [1], cloud storage [2], and cloud download [3].

As a typical cloud application, the cloud download arises, such as QQ Cyclone Download of Tencent company (http://xf.qq.com/), which has attracted millions of users and supports most of the mainstream content distribution protocols. In cloud download system [3], once a user specifies his resource need by sending request, the system downloads the file from a couple of Internet content providers, stores it in the cloud cache for a period, and notifies the user. Then, during this period, the user can retrieve his requested file from the system with higher speed in any place at any time. The other outstanding advantage of cloud download is that the user can be offline after requesting instead of keeping online in traditional download system. Due to the cloud download, the user saves time and energy. Especially, there is remarkable save when the requested file has low speed from Internet.

In cloud download, the time cost of user is determined by how long the user was notified and how long he took on retrieving the file, which are respectively called as the response time and the retrieving time. The combination of these times can be used to measure the quality of service. More time cost of user results in lower user satisfaction, which will cause users to move away. For the service provider, it controls user satisfaction by scheduling when to process the file and when to notify the user, both of which influence the overhead of renting cloud storage resources to store the file in cloud cache. More cost of storage resource leads to smaller profit the service provider earns.

As a business provider in cloud, the service provider of cloud download also aims to satisfy users furthest with minimized cost of storage resource. If the cloud download system processes the request and notifies the user immediately once it received the request, the time cost of user is likely minimized. But the time cost of storing the file in cloud cache may be much large. Consequently, both costs of user and system cannot be minimized together. So, the cloud download system faces great challenge on

Received date: 09-12-2013
Corresponding author: XU Ying-ying, E-mail: yingyingxuchina@163.com
DOI: 10.1016/S1005-8885(14)60315-0

scheduling processing job and notification to meet the conflicting objectives, i.e., the provider profit maximization and user satisfaction optimization. This paper is devoted to finding the scheduling strategy to make tradeoffs.

The contribution in this article is: based on the utility theory in economics, an economic model is built to formulate tradeoffs between user satisfaction and system profit in cloud download. By solving the model, the optimal job and notification scheduling strategy is obtained. Then it is demonstrated that strategy without intelligent scheduling is not the best choice, e.g., system starts processing job and notifies a user immediately after he requests. In comparison, our strategy can obtain higher user satisfaction under certain service profit target, or higher provider profit with certain user satisfaction bound.

The remainder is organized as follows. The related work is reviewed in Sect. 2. Sect. 3 describes the detailed download process in cloud download, gives the definition of the user satisfaction and the provider profit, and builds an economic model to explore the job and notification scheduling strategy making tradeoffs. The optimal strategies are proposed in several scenarios in Sect. 4. Then the performance of these strategies is discussed with real world cases in Sect. 5. Finally, Sect. 6 concludes the paper.

2 Related work

Study concerned about resources allocation while overlooking economic profit in cloud, which mostly considered minimized consumption of cloud resources, optimal performance for user, and enhanced system scalability. For example, work in Refs. [4–7] did research on service selection based on reputation mechanism, the optimal selection of data centers to minimize latency, fair allocation by a new cloud control mechanism, and a heavyweight holistic toolkit to cloud service provisioning for both the service and infrastructure providers.

Being a highly concerned issue, the resource allocation has some market-based methods in cloud, which can reveal the true needs of users and allocate resources more efficient. Li has proposed a flat rate pricing mechanism and formulated allocation to maximize the sum of the utility function of each user under the constraints of fairness [8]. In Ref. [9], Young developed pricing model and proposed profit-driven strategies for scheduling requests on service instances. A hierarchical auction model has been introduced to allocate network resources [10]. Recently, Chard has reduced the failure and reallocation rate by addressing limitations of economic allocation adoptions [11]. An analytical model has been developed to solve user-side resources allocations [12].

However, only a few work was reported for scheduling strategy in cloud with respect to a cloud application how to schedule the resource to make tradeoffs between its profit and user satisfaction, and much less suit cloud download. In Ref. [13], the problem is addressed for service provider that how to rent a set of virtual machine (VM) instances from an infrastructure service provider to make tradeoffs between its profit and user satisfaction.

The authors's work differs from previous studies in following aspects: First, satisfaction of users is defined as a mathematic function related to the response time, the retrieving time and the service price, which is mainly different from others. Second, the job and notification scheduling strategy determines when to process the requested file and when to notify the user respectively, which will determine the lengths of the response time and the retrieving time of user. Finally, since expansion of cached files in cloud download, the proposed strategy is a lightweight approach to implement easily to improve utilization of cloud storage resources.

3 Economic model

After analysis of a download process in cloud download, the satisfaction of users and the profit of service providers are investigated respectively, and their definitions are given. The economic model is subsequently proposed to formulate the challenge on how to schedule job and notification.

3.1 System architecture

The system architecture of cloud download is illustrated in Fig. 1.

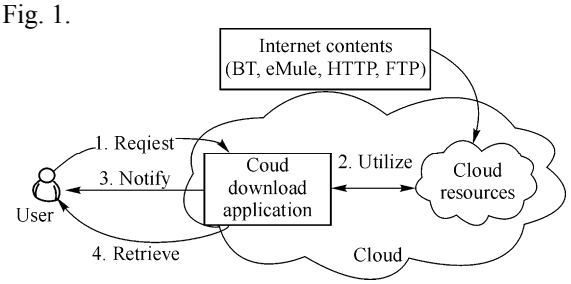

Fig. 1 The system architecture of cloud download

All the bold arrows depict exchange of file content data, and the process is not always carried out in order of the