

클라우드 컴퓨팅에서 데이터 분산 최적화를 위한 방법에 대한 연구

밤복흥, 이슬람, 고메즈 마우리시오, 모하마드 아 잠, 허의남
경희대학교 컴퓨터공학

e-mail : hung205a2@yahoo.com, {motahar, mgomez, aazam, johnhuh}@khu.ac.kr

A Novel Approach for Optimizing Data Distribution in Cloud Computing

Pham Phuoc Hung, Md. Motaharul Islam, Mauricio A. G. Morales, Mohammad Aazam, Eui-Nam Huh
Dept. of Computer Engineering, Kyung Hee University

Abstract

Modern day despite technology advancements that manufacture a new generation of mobile devices with generous resources, the fact that they can offer only limited processing capacity still remains a painful experience. So far, a number of research studies have been carried out, trying to eliminate problems arising from shortcomings in the connection between thin clients and cloud networks, yet little have been found efficient. In this paper, we present a novel approach, taking advantage of collaboration of thin and thick clients, particularly aiming at optimizing data distribution by splitting data and utilizing cloud computing (CC) resources so that expected Quality-of-Service (QoS) requirements can be met. Moreover, we conduct simulations to evaluate our approach. Our results evaluation shows that our approach has better performance than existing approaches.

1. Introduction

With worldwide shipments of smartphones (487.7 million) exceeding PCs (414.6 million including tablets) in 2011 [1], and in the US alone, more users predicted to access the Internet from mobile devices than from PCs by 2015 [2], we have witnessed a stunning success of mobile computing, which reflects a significant change in users' computer and internet usage. Mobile Computing, despite the amazing convenience and flexibility it brings, has its weakness in the ability to perform heavy computing tasks or high data transferring due to limitation in mobile device's resources (CPU, memory or battery life) and/or in the limited network bandwidth [3]. This problem, fortunately, has been made less severe thanks to the recent adoption of mobile cloud computing (MCC), which provides mobile applications the capacities of cloud servers and storage together with the benefits of mobile devices and mobile connectivity or collaborating with external devices to get more resources.

So far, a number of research studies have been carried out, trying to eliminate problems arising from shortcomings in the connection between thin clients and cloud networks, yet little have been found efficient. In [4], for example, Gonzalo Huerta-Canepa presents guidelines for a framework to create virtual mobile cloud computing providers. This framework takes advantage of nearby thin clients to create on-the-fly connection, thus avoiding the need to connect to infrastructure-based clouds. The downside of this method, however, lies in the thin clients' capacity restriction and low bandwidth between them and cloud. Thick clients like laptop or desktop computers, on the other hand, normally own greater capacities and connections with much higher

bandwidth. Good bandwidth is very important, the higher bandwidth we have the higher quality of services we get [5]. It is therefore recommended thin clients get connected with thick clients so that the strengths of the latter can help relieve the weaknesses of the former.

In this paper, we introduce a new architecture that collaborates thin clients and thick clients, accordingly enhancing the former's capacities. Furthermore, we propose a strategy to optimize data distribution, especially big data in cloud computing. Besides, we pick out algorithms to perform resource allocation by using optimization model to calculate the minimum number of thick clients so as to meet the quality of service. To evaluate our proposed approach, we conduct a number of simulations whose results show that collaboration of thin-thick clients associated with splitting big data to match the different thick clients' relay capacity is essentially achieve better performance than other existing solutions.

Organization of the paper is as follows: The second section provides information about related work on previous approaches solving some part of the problems mentioned above. Section 3 details the motivating scenario which explains why we have come up with the proposed approach. Section 4 presents our proposed architecture while section 5 reports the insight of the implementation, performance evaluation and discussion. The last section concludes the paper and future work.

2. Related Study

There have been numerous studies which attempt to solve some parts of above problems. In [6], authors propose a new

approach for efficient cloud-based synchronization of a number of distributed file system hierarchies. They use a master-slave architecture for propagation of data to reduce traffic. In [7], researchers demonstrate that some resource scheduling techniques can be effective in mitigating the impacts that negatively influence application response time and system utilization. N. M. Andreolini [8] and P. Fan [9] introduce the impact of the data transfer delay on the performance but they do not reckon to use bandwidth efficiently. Gueyoung Jung, Nathan [10] present a way to make a parallel process to a big data to increase performance in Federated Clouds but they do not consider how much resources should be used.

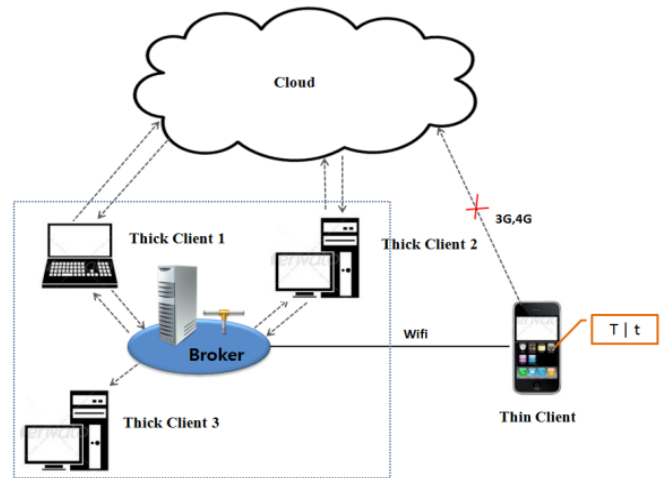
Similar to our approach, other research efforts have been made to integrate mobile devices and cloud computing. In [11], X. Luo suggests a new idea of using cloud to improve mobile device's capability. Marinelly [12] innovates Hyrax, which allows mobile devices to use cloud computing platforms. The researcher introduces the idea of using mobile devices as resource providers, but experiment is not integrated. In [13], the authors just concentrate on using partition policies to hold the effect of application on mobile devices, but they do not solve any other matter related to mobile cloud computing. Also, Zhong and Longzhao [14] discuss the integration of CC into mobile internet and are able to exemplify their arguments with typical successful business models in the market yet without particular performance benchmarking.

3. Motivating Scenario

The following scenario reflects the essential benefits behind the proposal of the thin-thick client collaboration. A mobile user wants to watch online movies using her mobile phone. At first, he/she intends to play a selected movie using the cellular network (e.g. 3G). Unfortunately, in the area he/she is living there is poor 3G coverage, which results in very low network connectivity and thus it is almost impossible for him/her to enjoy movies. Moreover, he/she recognizes that the choice to download a movie using the cellular network connection is too expensive. So he/she tries to find an alternative way to access the Internet by having his/her phone (known as thin client) connected to her workplace Wi-Fi network, which includes a group of laptop and desktop computers (also known as thick clients) as illustrated in Figure 1.

We make an assumption that there is a broker who is managing several thick clients; each can access a provider. We further assume that the thin client can send a request (e.g., movie streaming) to the broker who chooses the optimal number of thick clients to serve this request. Each of the chosen thick clients will establish a connection to the corresponding cloud provider in which thin client desires to access. Once a request is received, the cloud provider will process it and reply to the connecting thick client with suitable content (i.e., part of the requested movie) through these chosen thick clients' paths. Because the bandwidth between the thick clients and cloud networks can be very high (e.g., DSL or cable connections), the parts of data can be delivered to thick clients in relatively short time before they are collected at the broker computer, where they are then combined and forwarded to the requesting thin client.

Thereby, the thin client takes advantage of multiple thick clients' relay to enhance the data distribution from cloud networks, consequently enhancing its computing capabilities.



(Figure 1) Scenario of requesting a cloud service

The above scenario shows the potential benefit of utilizing the joint work between thin clients and thick clients in a typical mobile cloud computing environment. Such collaboration promisingly increases the opportunity of using resources efficiently. With that in mind and with CC platform stays ready and emerging on the market, our paper aims at architecting a network design based on the thin-thick client collaboration, attempting to address the following issues:

- Optimizing the data distribution between mobile thin client and cloud network.
- Picking the best resource allocation strategy among available ones to determine the minimum number of thick clients to satisfy the thin client's QoS requirement.

4. System Architecture

In this section, we are going to describe our system architecture to solve the above problems. Figure 3 illustrates the general design of the proposed system. Different from other approaches that follow a $1/m/1$ model which states that one thin client is served by one server through multiple-paths, our system can be viewed as an $1/m/1/m$ model. Our system model states that from a cloud provider, data is divided into multiple chunks and those chunks are transferred to multi thick clients (processors). After receiving the data, they combine the data into one at a broker then send it to destinations. For ease of presentation, our system can be divided into the following two layers as shown in Figure 2.

Layer 1:

This layer involves: Splitting and distributing data from cloud provider to processors with different capacities according to bandwidth of Internet connection.

For clarity, we give the important definition and assumption for our system. First, split each data from the cloud provider into chunks $\{ch_1, ch_2, \dots, ch_n\}$ with different sizes depending on bandwidth. $w(ch_i)$ is the size of a chunk ch_i . b_i is bandwidth from a VM to a processor. Therefore, time spent for the transferring a chunk ch_i from VMs to a processor is $w(ch_i)/b_i$. For parallelization, the time to transfer chunks to

processors should be equal.

$$\frac{w(ch_1)}{b_1} = \frac{w(ch_2)}{b_2} = \dots = \frac{w(ch_i)}{b_i} = t$$

$$\text{Set } S = w(\text{data}) = \sum_{i=0}^n w(ch_i) = t \sum_{i=0}^n b_i$$

$$\text{Thus, } w(ch_i) = t * b_i = \frac{S}{\sum_{i=0}^n b_i} * b_i$$

According to this value, we can determine the size of each chunk to adapt with bandwidth. Next step is to sort the processors $\{p_1, p_2, \dots, p_p\}$ depending on their capacities. The processor which has higher capacity will receive big chunks and the processor which has lower capacity will receive small chunks.

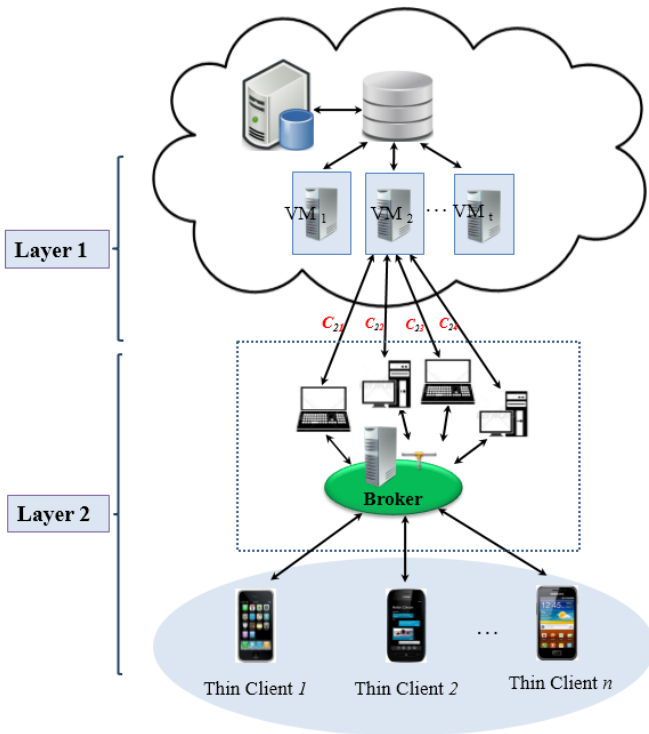


Figure 2. Layering architecture of the proposal

Layer 2:

In this layer we consider the following: (1) calculating the minimum number of thick clients to satisfy QoS and (2) combining the data then transfer it to a thin client.

a) Calculating the minimum number of thick clients to satisfy QoS.

First, we define terms used in this part. Then we introduce an optimization model that formalizes the problem based on the response time and workload of thick clients.

Suppose that we are given n thin clients that are to be served by M thick clients. Each thick client has a certain workload W_m and response time T_m^R . The corresponding workload as well as response time of each thin client is W_n and T_n^R .

Considering multiple types of thick clients, such as laptop, personal computer, the optimization problem can be formulated as following:

Minimize M

$$a_m \in \{0,1\}$$

subject to

$$\sum_{n \in S_l^m} W_n < \sum_{m=1}^M a_m W_m,$$

$$\text{Min} \{T_n^R, \forall n \in S_l^m\} > T_m^R, \forall m.$$

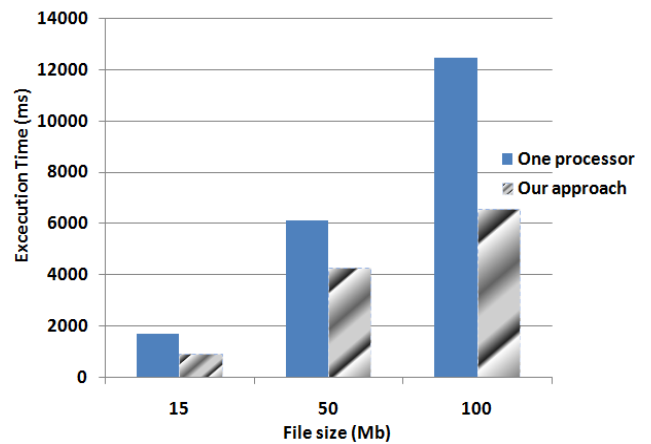
Our objective is to minimize the number of thick clients to satisfy variety of thin clients which connect to broker, while first set of constraints makes sure that total workload of thin clients does not exceed the total workload of thick clients. The second set of constraints ensures that the response time of each thick client is less than the response time required by thin clients. By using this optimization model, it is easy for us to calculate the minimum number of thick clients to guarantee the QoS.

b) Combining the data then transfer it to a thin client:

After data from cloud service has been received, instead of using peer-to-peer synchronization between all processors, which might make communications more complex, we consider that a processor (the broker computer) acts as a master which receives data from others processors to decrease the complexity due to firewall between processors.

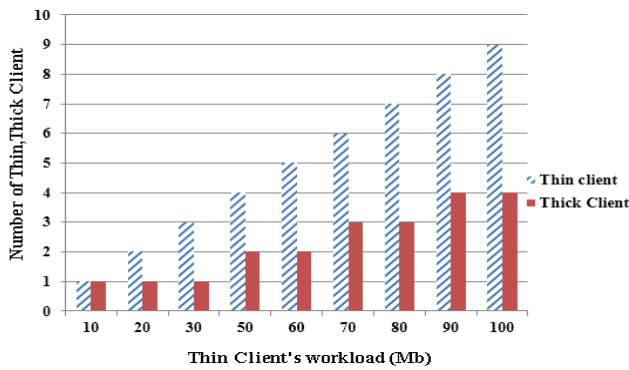
5. Implementation

In this section, we use numerical simulations to evaluate compare our approach's performance with others'. The simulations is developed in Java with jdk-7u7-i586 and Netbeans-7.2. We compare the processing time of transferring big data from a source to a destination of our system and other approach which has one processor receiving data. Through experiment result in Figure 3, it is obvious that our approach leads to a better performance.



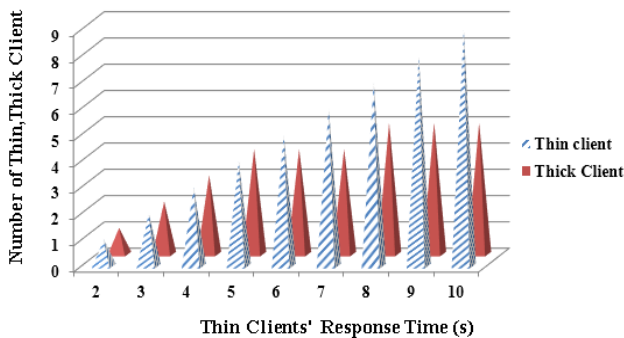
(Figure 3) Comparison our approach with other approach.

We also conducted simulations to calculate the minimum number of thick clients which are able to satisfy thin clients' requirement with different workload and response time. In Figure 4, when the workload of thin clients increases, the number of thick clients also increases.



(Figure 4) Thin thick clients with different workload.

The same result happens once response time of thin clients decreases like Figure 5. Especially, the more thin clients in the area with short response time there are, the more thick clients should be used to guarantee the QoS.



(Figure 5) Thin thick clients with different response time.

6. Conclusion

In this paper, we have introduced a network architecture in mobile cloud computing that promotes the joint work of thin and thick clients to enhance thin client capacities. In particular, our work focuses on optimizing the data distribution from cloud networks to mobile clients and has supported decision making on the minimum number of thick clients should be used to guarantee QoS. Moreover, we have conducted simulations to evaluate our method. As can be observed from the implementation, our proposed approach can improve resource allocation efficiency and has better performance than other existing approaches.

In the future, we will amend and extend our approach in a variety of circumstances to achieve higher reliability and better performance of the proposal.

Acknowledgement

This research was supported by Next-Generation Information Computing Development Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (2012- 0006418). The corresponding author is Eui-Nam Huh.

References

- [1] Canals, "Smart phones overtake client PCs in 2011", <http://www.canalys.com/newsroom/smart-phones-overtake-client-pcs-2011> ", Feb. 2012.
- [2] IDC, "More Mobile Internet Users Than Wireline Users in the U.S. by 2015", <http://www.idc.com/getdoc.jsp?containerId=prUS23028711> , Sep. 2011.
- [3] Pham Phuoc Hung, Bui Tuan-Anh, Eui-Nam Huh, "A Solution of Thin-Thick Client Collaboration for Data Distribution and Resource Allocation in Cloud Computing ", ICOIN 2013.
- [4] Gonzalo Huerta-Canepa, "A Virtual Cloud Computing Provider for Mobile Devices", MCS'10, June 15, 2010, San Francisco, California, USA.
- [5] Tien-Dung Nguyen, "Service Image Placement for Thin Client in Mobile Cloud Computing", 2012 IEEE, DOI 10.1109/CLOUD.2012.39.
- [6] SandeshUppoor, Michail D. Flouris, and AngelosBilas, "Cloud-based Synchronization of Distributed File System Hierarchies", Cluster Computing Workshops and Posters, 2010 IEEE International Conference , pp 1 - 4.
- [7] Javier Delgado, Liana Fong, "Efficiency Assessment of Parallel Workloads on Virtualized Resources" , 2011 Fourth IEEE International Conference.
- [8] Pei Fan, JiWang,ZibinZheng, Michael R. Lyu, "Toward optimal deployment of communication-intensive cloud applications," Cloud Computing (CLOUD), 2011 IEEE International Conference,4-9 July 2011,pp 460 – 467.
- [9] Michael Kwok, "Performance Analysis of Distributed Virtual Environments", PhD Thesis, University of Waterloo, Ontario, Canada, 2006.
- [10] Gueyoung Jung, Nathan, "Synchronous Parallel Processing of Big-Data Analytics Services to Optimize Performance in Federated Clouds", 2012 IEEE, USA.
- [11] X. Luo, "From Augmented Reality to Augmented Computing: A Look at Cloud-Mobile Convergence," International Symposium on Ubiquitous Virtual Reality, 8-11 July 2009, pp 29 - 32..
- [12] E. Marinelli, "Hyrax: Cloud Computing on Mobile Devices using MapReduce," Master Thesis Draft, Computer Science Dept., CMU, September 2009.
- [13] Giurgiu, O. Riva, D. Juric, I. Krivulev, and G. Alonso, "Calling the cloud: enabling mobile phones as interfaces to cloud applications," International Conference on Middleware, New York, NY, USA, 2009, pp. 1-20.
- [14] Zhong, Longzhao, "Cloud computing applied in the mobile Internet", 2012 7th International Conference on Computer Science & Education (ICCSE), July 2012