IOSR Journal of Engineering (IOSR JEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 05-08

Optimal Service Providing In Cloud Computing With Performance Guarantee

Elamathi.V[1], Dr. PriyaRadhikadevi .T[2]
[1](PG Scholar, Dept of computer science and Engineering, Mailam Engineering College.)
[2](.Head of department ,Dept of computer science and Engineering, Mailam EngineeringCollege)

Abstract: *As an effective and efficient way to provide computing resources and services to customers on demand, cloud computing has become more and more popular. From cloud service providers' perspective, profit is one of the most important considerations, and it is mainly determined by the configuration of a cloud service platform under given market demand. However, a single long-term renting scheme is usually adopted to configure a cloud platform, which cannot guarantee the service quality but leads to serious resource waste. In this paper, a double resource renting scheme is designed firstly in which short-term renting and long-term renting are combined aiming at the existing issues. This double renting scheme can effectively guarantee the quality of service of all requests and reduce the resource waste greatly. Secondly, a service system is considered as an M/M/m+D queuing model and the performance indicators that affect the profit of our double renting scheme are analyzed, e.g., the average charge, the ratio of requests that need temporary servers, and so forth. Thirdly, a profit maximization problem is formulated for the double renting scheme and the optimized configuration of a cloud platform is obtained by solving the profit maximization problem. Finally, a series of calculations are conducted to compare the profit of our proposed scheme with that of the single renting scheme. The results show that our scheme can not only guarantee the service quality of all requests, but also obtain more profit than the latter.*

I. Introduction

As an effective and efficient way to consolidate computing resources and computing services, clouding computing has become more and more popular [1]. Cloud computing centralizes management of resources and services, and delivers hosted services over the Internet. The hardware, software, databases, information, and all resources are concentrated and provided to consumers on-demand [2]. Cloud computing turns information technology into ordinary commodities and utilities by the the pay-per-use pricing model [3, 4, 5]. In a cloud computing environment, there are always three tiers, i.e., infrastructure providers, services providers, and customers (see Fig. 1 and its elaboration in Section 3.1). An infrastructure provider maintains the basic hardware and software facilities. A service provider rents resources from the infrastructure providers and provides services to customers. A customer submits its request to a service provider and pays for it based on the amount and the quality of the provided service [6]. In this paper, we aim at researching the multiserver configuration of a service provider such that its profit is maximized. Like all business, the profit of a service provider in cloud computing is related to two parts, which are the cost and the revenue. For a service provider, the cost is the renting cost paid to the infrastructure providers plus the electricity cost caused by energy consumption, and the revenue is the service charge to customers. In general, a service provider rents a certain number of servers from the infrastructure providers and builds different multiserver systems for different application domains. Each multiserver system is to execute a special type of service requests and applications. Hence, the renting cost is proportional to the number of servers in a multiserver system [2]. The power consumption of a multiserver system is linearly proportional to the number of servers and the server utilization, and to the square of execution speed [7, 8]. The revenue of a service provider is related to the amount of service and the quality of service. To summarize, the profit of a service provider is mainly determined by the configuration of its service platform.

To configure a cloud service platform, a service provider usually adopts a single renting scheme. That's to say, the servers in the service system are all long-term rented. Because of the limited number of servers, some of the incoming service requests cannot be processed immediately. So they are first inserted into a queue until they can handled by any available server. However, the waiting time of the service requests cannot be too long. In order to satisfy quality-of-service requirements, the waiting time of each incoming service request should be limited within a certain range, which is determined by a service-level agreement (SLA). If the quality of service is guaranteed, the service is fully charged, otherwise, the service provider serves the request for free as a penalty of low quality. To obtain higher revenue, a service provider should rent more servers from the infrastructure providers or scale up the server execution speed to ensure that more service requests are processed with high

service quality. However, doing this would lead to sharp increase of the renting cost or the electricity cost. Such increased cost may counterweight the gain from penalty reduction. In conclusion, the single renting scheme is not a good scheme for service providers. In this paper, we propose a novel renting scheme for service providers, which not only can satisfy quality-of-service requirements, but also can obtain more profit. Our contributions in this paper can be summarized as follows.

- A novel double renting scheme is proposed for service providers. It combines long-term renting with short-term renting, which can not only satisfy quality-of-service requirements under the varying system workload, but also reduce the resource waste greatly.
- A multiserver system adopted in our paper is modeled as an M/M/m+D queuing model and the performance indicators are analyzed such as the average service charge, the ratio of requests that need shortterm servers, and so forth.
- The optimal configuration problem of service providers for profit maximization is formulated and two kinds of optimal solutions, i.e., the ideal solutions and the actual solutions, are obtained respectively.
- A series of comparisons are given to verify the performance of our scheme. The results show that the proposed Double-Quality-Guaranteed (DQG) renting scheme can achieve more profit than the compared Single-Quality-Unguaranteed (SQU) renting scheme in the premise of guaranteeing the service quality completely.

The rest of the paper is organized as follows. Section 2 reviews the related work on profit aware problem in cloud computing. Section 3 presents the used models, including the three-tier cloud computing model, the multiserver system model, the revenue and cost models. Section 4 proposes our DQG renting scheme and formulates the profit optimization problem. Section 5 introduces the methods of finding the optimal solutions for the profit optimization problem in two scenarios. Section 6 demonstrates the performance of the proposed scheme through comparison with the traditional SQU renting scheme. Finally, Section 7 concludes the work.

II. Existing System

- In general, a service provider rents a certain number of servers from the infrastructure providers and builds different multi-server systems for different application domains. Each multiserver system is to execute a special type of service requests and applications. Hence, the renting cost is proportional to the number of servers in a multiserver system. The power consumption of a multiserver system is linearly proportional to the number of servers and the server utilization, and to the square of execution speed. The revenue of a service provider is related to the amount of service and the quality of service. To summarize, the profit of a service provider is mainly determined by the configuration of its service platform.
- To configure a cloud service platform, a service provider usually adopts a single renting scheme. That's to say, the servers in the service system are all long-term rented. Because of the limited number of servers, some of the incoming service requests cannot be processed immediately. So they are first inserted into a queue until they can handle by any available server.

Disadvantages Of Existing System:

- The waiting time of the service requests is too long.
- Sharp increase of the renting cost or the electricity cost. Such increased cost may counterweight the gain from penalty reduction. In conclusion, the single renting scheme is not a good scheme for service providers.

III. Proposed & Modification System

- ❖ In this paper, we propose a novel renting scheme for service providers, which not only can satisfy quality-of-service requirements, but also can obtain more profit.
- ❖ A novel double renting scheme is proposed for service providers. It combines long-term renting with short-term renting, which can not only satisfy quality-of-service requirements under the varying system workload, but also reduce the resource waste greatly.
- ❖ A multiserver system adopted in our paper is modeled as an M/M/m+D queuing model and the performance indicators are analyzed such as the average service charge, the ratio of requests that need shortterm servers, and so forth.
- ❖ The optimal configuration problem of service providers for profit maximization is formulated and two kinds of optimal solutions, i.e., the ideal solutions and the actual solutions, are obtained respectively.
- ❖ A series of comparisons are given to verify the performance of our scheme. The results show that the proposed Double-Quality-Guaranteed (DQG) renting scheme can achieve more profit than the compared

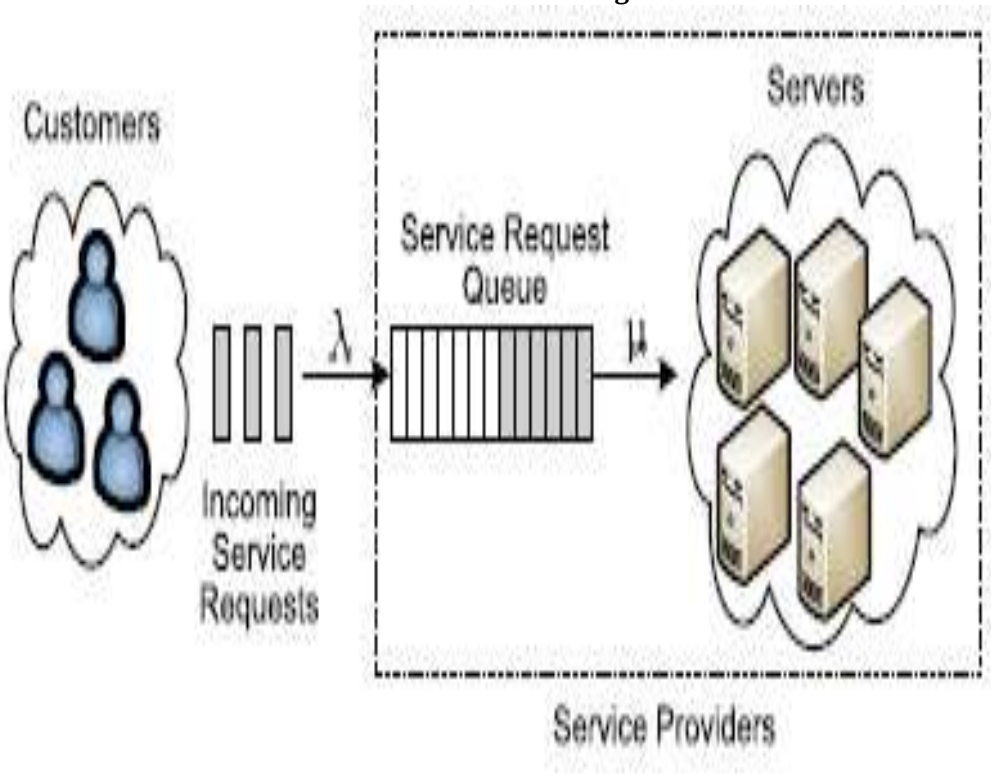
Single-Quality-Unguaranteed (SQU) renting scheme in the premise of guaranteeing the service quality completely.

Advantage Of Proposed System

- ✓ Since the requests with waiting time D are all assigned to temporary servers, it is apparent that all service requests can guarantee their deadline and are charged based on the workload according to the SLA. Hence, the revenue of the service provider increases.
- ✓ Increase in the quality of service requests and maximize the profit of service providers.

This scheme combines short-term renting with long-term renting, which can reduce the resource waste greatly and adapt to the dynamical demand of computing capacity.

IV. Architecture Diagram



V. Conclusion

written stories to a specific low-level KR. This new framework is able to reason with uncertainty, to integrate training from annotated data and constraints encoding information on mutually exclusive values, beyond evidence from external sources, such as information from the language model [20]. Similar to other methods for structured prediction, the mapping aims at predicting the most likely structure by searching in the large search space derived from the exponential explosion of instance combinations, i.e. MAP inference. Therefore, an algorithm based on GA, able to exploit some properties of the Bayesian network, see (15) and (6), was developed for the statistical inference, requiring less CPU time than state-of-the-art tools, while providing parallel scalability to deal with larger domains. Moreover, the new constrained learning algorithm for Bayesian networks yielded performance gains in predicting the most likely structure given new sentences (unseen during the training).

References

[1]. K. Hwang, J. Dongarra, and G. C. Fox, Distributed and Cloud Computing. Elsevier/Morgan Kaufmann, 2015.
[2]. J. Cao, K. Hwang, K. Li, and A. Y. Zomaya, "Optimal multiserver configuration for profit maximization in cloud computing," IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 6, pp. 1087–1096, 2016.
[3]. A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, and I. Stoica, "Above the clouds: A berkeley view of cloud computing," Dept. Electrical Eng. and Comput. Sciences, vol. 28, 2017.
[4]. R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud computing and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility," Future Gener. Comp. Sy., vol. 25, no. 6, pp. 599– 616, 2016.

[5]. P. Mell and T. Grance, "The NIST definition of cloud computing. national institute of standards and technology," Information Technology Laboratory, vol. 15, p. 2015.
[6]. J. Chen, C. Wang, B. B. Zhou, L. Sun, Y. C. Lee, and A. Y. Zomaya, "Tradeoffs between profit and customer satisfaction for service provisioning in the cloud," in Proc. 20th Int'l Symp. High Performance Distributed Computing. ACM, 2017, pp. 229–238.